

Ridwan Maulana
Michelle Helms-Lorenz
Robert M. Klassen *Editors*

Effective Teaching Around the World

Theoretical, Empirical, Methodological
and Practical Insights

OPEN ACCESS

 Springer

Effective Teaching Around the World


Ridwan Maulana • Michelle Helms-Lorenz
Robert M. Klassen
Editors


Effective Teaching Around the World


Theoretical, Empirical, Methodological
and Practical Insights

 Springer

Editors

Ridwan Maulana 
Department of Teacher Education
University of Groningen
Groningen, The Netherlands

Michelle Helms-Lorenz 
Department of Teacher Education
University of Groningen
Groningen, The Netherlands

Robert M. Klassen 
Department of Education
University of York
York, UK



ISBN 978-3-031-31677-7 ISBN 978-3-031-31678-4 (eBook)
<https://doi.org/10.1007/978-3-031-31678-4>

The publication of this book was funded partially by the Dutch Research Council (NWO) with grant number 36.201.068; University of Groningen; and by the majority of the contributors' institutions.

© The Editor(s) (if applicable) and The Author(s) 2023. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

This work is dedicated to researchers, practitioners, policy makers, all colleagues and friends trying to make a difference in education by seeking to understand the complex nature of teaching and improve the quality of teaching around the world.

Foreword

We have all experienced two years of the largest (unplanned) educational experiment in our lifetimes. Schools as we knew them were closed, and distance or hybrid learning was introduced. There was equity, resourcing, death, unemployment, and many homes were not safe havens or ideal learning places for children. Many parents soon realized that they did not have the skills of teachers to motivate, sustain, and teach their children. There are already two meta-analyses published on the effects of COVID showing minimal losses in the trajectory of learning from the start till the end of the year (Konig & Frey, 2022; Zierer, 2021). Compared to the usual gains made, on average, the gains during COVID were minimally lower compared to the previous 10 years.

This minimal change, surely, attests to the effectiveness of educators – who worked so hard to ensure there was no learning loss, that the gains typically made over a year were (almost) maintained, and that students were as minimally disadvantaged as possible (Hattie, 2021). The greatest travesty of COVID schooling is rushing back to the old normal and not pausing to learn about what was so effective during COVID teaching to augment our older grammar of schooling. In the old grammar of schooling, teachers talk a lot (80-90%), ask 100-150 questions a day requiring less than three-word answers about the facts, and too many students come to class to watch teachers work. It is not possible in COVID teaching to replicate this, as teachers moved from in-front control to triage, from talking to listening, to (gradually) releasing their responsibility, and teaching students to become their own teachers and work effectively efficiently with their peers.

This book is thus timely as it aims to detect the greatest efficacy in our knowledge of teaching, and if only we could then de-implement that which does not feature and augment the effectiveness with learnings from COVID teaching, we could serve more students, entice more to love the learning at school, focus on progress to achievement, and teach the optimal strategies of learning. The book illustrates the richness of exemplary practice in our schools, and if only we could learn to scale up this quality, then so much the better for all students.

I see effective teaching in terms of maximizing the impact on the learning, achievement, and well-being of all students, such that students become their own teachers, learn how to learn alone and with others, and more fully appreciate the importance of precious knowing and understanding the world they will create for us all. Effectiveness is in terms of impact (which begs the moral questions: impact about what, for whom, and how large is this impact) not in terms of specific correlates, methods, or personal attributes. Throughout this book, the answer to this question about effective teaching is not straightforward, varies depending on context and where the student is in the learning cycle, and the authors have taken on a monumental task to tease through these issues.

The chapters outline the many models, but as is so common in our discipline, there are few empirical or theoretical comparisons of these models. Of course, there are exceptions and these are noted. For example, the Dynamic Approach to Teaching Improvement is one of the more powerful models, and most important it can be used to promote improvement in effectiveness. The five principles are well-evidenced in the research and underline much of the queries in the remainder of the book. Their terms are frequency, focus, stage, quality, and differentiation, and I translate these into: how much dosage of a teaching method is needed to get impact, is it aimed at knowing that, how, or with (surface deep, transfer), where in the learning cycle is the student (early exposure, consolidating, relational, extending), the fidelity to the method, and the extent of adaption (and too much adaption can be the killer of effectiveness). I particularly note that ‘differentiation’ does not mean different activities for different groups of students, but more allowing for different times and ways of progression towards appropriately challenging success criteria. This seems not always agreed in other chapters and differentiation remains a fuzzy concept. For example, Rubie-Davies (2010) showed that high impact teachers rarely mention differentiation, as they are averse to different activities for different groups, preferring to allow different pathways and different times to all their students. Similarly, all students deserve a learning intervention plan, need to be taught to become assessment capable to learn about their own progress, and given feedback that helps them know where to move next.

Throughout the book, there are so many factors cited as critical to effectiveness, although there are many common denominators. But it is the constructive alignment of these factors with the level of cognitive complexity that is critical. It is how teachers differentiate (to use that word again in a different way) their teaching methods to the learning cycle, and most critical have multiple teaching methods as if the first does not work they have alternates to use in re-teaching.

The reality of implementation is often the killer of great empirical models, and more attention to dosage, fidelity, quality, and adaptation is needed. Similarly, grounding models of effectiveness in exemplary teachers practice is important (many an academic may say it works in practice but may not work in theory!). van Geel et al. provide an excellent demonstration of the importance of focusing on implementation. When comparing Differentiated Instruction and Assessment for Learning, they note that AfL emphasizes eliciting evidence during the lesson, and

DI emphasizes pro-active alignment of instruction and activities based on students' needs. Similar factors but different emphases.

Other models focus on motivating students, although there are few students I know who do not come to class with deep wells of motivation, but maybe not to spend these resources on school subjects. It is more why do this rather than that, and not how to push or pull students into a lesson (Hattie et al., 2020).

All this requires major cognitive demands on teachers, especially new teachers who are often thrust into classrooms with the same demands as more experienced peers. After reading this book, there is a sense of marvel at the depth of cognitive complexity demanded from today's teachers. Johansson provides worrying data about the drop in academic prerequisites to become a teacher: a massive 25% drop in GPA in grade 9 for new teachers from 1996 to 2016; and an increase from 15% to 26% in non-certified teachers in schools. Surely this is going in the wrong direction. There is a threat to the school system if we do not recognize the cognitive as well as personal and emotional demands and ensure we start with the most optimal cohort of students in initial teacher education programs. The increase of amateurs in schools should be the most worrying dilemma of schools in well-resourced countries. Expertise is expensive, worth fighting for, and is the essence of our profession (Rickards et al., 2021).

There is richness in the many quantitative and qualitative methods to identify effectiveness, and many chapters show the value of these methods across countries, curricula, and age levels. Often missed are student perspectives of effective teaching. A valued contribution is the chapter by Bijlsma and Röhl showing how student evaluations of the impact of their teachers can help triangulate other information on effectiveness. Perhaps the next major breakthrough in methods is automating classroom observation methods. In our own VisibleClassroom project, teachers turn on an app on their iPhone, teach the lesson, and immediately retrieve a transcript of their lesson and a report (which uses AI) to review 18 dimensions of effective teaching. Since we commenced, others are making critical AI advances to analyze the observations, and access to these reports and interpretation will accelerate our evidence of impact (Liu & Cohen, 2021).

Many chapters delve into this richness of comparing the notions and implications of models of effectiveness across countries and cultures. I recall working with a colleague comparing teacher excellence in China and NZ, and she claimed there was little difference. But delving deeper, she noted that in China it was normal for the head teacher to teach a model class and then for the staff to critique it – unheard of (almost) in Western schools. There is a culture of autonomy meaning each teacher can teach their way and dare there be critique of one's autonomy. We have much to learn how to make less the evidence of teaching less private impact, how to create safe and high trust staffrooms to have critique and debate about effectiveness, and how to elaborate each other's expectations, interpretations, and quality of evidence of impact. It is fascinating to see so many non-western countries investing in teacher quality, developing teacher standards, and seeking a robust manner to so do. In the West, we seem to love the politics of distraction and invest in buildings, curricula, and testing and minimize investment in expertise and standards.

This debate about effective teaching around the world will continue, and long may it but at the forefront of our research and practice. This ‘one-stop book’ goes a long way to advancing, promoting, and informing the debate, and there is indeed a richness herein.

Laureate Professor, University of Melbourne,
Author of Visible Learning Book
Parkville, VIC, Australia

John Hattie

References

- Hattie, J. A. C. (2021). *An Ode to expertise*. Corwin. <https://protect-au.mimecast.com/s/43HsCXLKNwFX9YWqYT62nlB?domain=f.hubspotusercontent10.net>
- Hattie, J. A. C., Hoddiss, F., & Kang, S. (2020). Theories of motivation: Integration and ways forward. *Contemporary Educational Psychology*, *61*, 101865.
- König, C., & Frey, A. (2022). The impact of COVID-19-related school closures on student achievement—A meta-analysis. *Educational Measurement: Issues and Practice*, *41*(1), 16–22.
- Liu, J., & Cohen, J. (2021). Measuring teaching practices at scale: A novel application of text-as-data methods. *Educational Evaluation and Policy Analysis*, *43*(4), 587–614.
- Rickards, F., Hattie, J. A. C., & Reid, C. (2021). *The turning point: Growing expertise, evaluative thinking, and the future of the teaching profession*. Routledge.
- Rubie-Davies, C. M. (2010). Teacher expectations and perceptions of student attributes: Is there a relationship?. *British Journal of Educational Psychology*, *80*(1), 121–135.
- Zierer, K. (2021). Effects of pandemic-related school closures on pupils’ performance and learning in selected countries: A rapid review. *Education Sciences*, *11*(6), 252.

Foreword

Because the topic of teaching effectiveness is of considerable importance and perennial interest internationally, it deservedly has been the focus of a vast amount of prior research and publications. In this comprehensive 36-chapter book, its editors (Ridwan Maulana, Michelle Helms-Lorenz and Rob Klassen) make an outstanding contribution by complementing, advancing and filling gaps in our knowledge about educational improvement and effective teaching.

As the book's title suggests, it encompasses insights that are theoretical, empirical, methodological and practical. These insights come from research and authors from many diverse countries (both more- and less-developed) and cultures. Audiences for the book include educational policy-makers, practitioners and researchers.

In many earlier publications, the work of teachers is regarded as being central and significant in students' learning. This volume is no exception.

An interesting and commendable inclusion is the book's closing chapter in which its three editors draw together insights, commonalities and differences across the book's many chapters, identify potential future research directions and, importantly, make recommendations for improving educational policy and practice in order that schools and teachers can better realise their educative potential.

The chapters' individual authors and the book's editors are to be congratulated on a significant, illuminating, scholarly and useful work on an internationally-relevant topic.

John Curtin Distinguished Professor, Curtin University
Perth, WA, Australia

Barry J. Fraser

Founding father of Learning Environments Research
and AERA Special Interest Group
Washington, DC, USA

Acknowledgements

The first international meeting on the International Comparative Analysis of Learning and Teaching (ICALT3/Differentiation) at the University of Groningen in 2016 brought together researchers, educators, practitioners, and policy makers from various partner countries to share knowledge and exchange ideas about teaching quality improvement. This meeting paved the way for continuous cooperation in effective teaching among key stakeholders in education in various countries across the five continents. This book is in part the result of this continuous international cooperation. We thank all committed partner countries within the Global Effective Teaching and Learning Network (GETLIN) and other external colleagues for contributing to this international book.

All the chapter contributions were initially assessed by the editors for suitability for the book. Chapters deemed suitable were sent to two reviewers to assess the scientific quality of the paper. All chapters included in this book were anonymously peer-reviewed twice. We are grateful to all reviewers for their constructive review. The editors were responsible for the final decision regarding acceptance or rejection of chapters. We are also indebted to Cor Suhre, Marjon Fokkens-Bruinsma, Lidewij van Katwijk, and Tim Huijgen from The University of Groningen for their help with reviewing.

We also would like to specifically thank to Bilge Gencoglu (University of Groningen) for providing editorial assistance during the process of preparing this book. The contribution of Sibel Telli (Canakkale University) in the beginning of the project and the assistance of Mirte de Vries (University of Groningen) are also acknowledged.

Finally, the biggest thanks goes to all research participants, mainly teachers and students, for participating in the research included in this book. Thank you to all for your excellent cooperation in co-creating insights towards understanding and improving the quality of education worldwide.

Contents

1	Prologue	1
	Ridwan Maulana, Michelle Helms-Lorenz, and Robert M. Klassen	
Part I Conceptualization and Measurement of Effective Teaching		
2	Using Educational Effectiveness Research for Promoting Quality of Teaching: The Dynamic Approach to Teacher and School Improvement	7
	Leonidas Kyriakides and Anastasia Panayiotou	
3	Teacher and Teaching Behaviour and Student Motivational Outcomes: Critical Reflections on the Knowledge Base and on Future Research	29
	Marie-Christine Opdenakker	
4	Teacher Recruitment in Sweden Over the Last Two Decades: How Has Entering Teachers' GPA Changed Over Time?	85
	Stefan Johansson	
5	Effective Teaching: Linking Outcomes of Active Citizenship to Learning Environments	97
	Gordon Sturrock and David Zandvliet	
6	Fostering Effective Teaching at Schools Through Measurements of Student Perceptions: Processes, Risks and Chances	119
	Hannah J. E. Bijlsma and Sebastian Röhl	
7	Differences in Perceived Instructional Quality of the Same Classrooms with Two Different Classroom Observation Instruments in China: Lessons Learned from Qualitative Analysis of Four Lessons Using TEACH and ICALT	137
	Jieyan Celia Lei, Zhijun Chen, and James Ko	

8	Measuring Teaching Skill of South Korean Teachers in Secondary Education: Detecting a Teacher's Potential Zone of Proximal Development Using the Rasch Model	165
	Wim van de Grift, Okhwa Lee, and Seyeoung Chun	
Part II Effective Teaching: Insights from Specific Countries		
9	Dialogic Interactions in Higher Vocational Learning Environments in Mainland China: Evidence Relating to the Effectiveness of Varied Teaching Strategies and Students' Learning Engagement	207
	Yanmin Zhao, Marc Kleinknecht, and James Ko	
10	Teaching Quality in Indonesia: What Needs to Be Improved?	225
	Yulia Irnidayanti and Nurul Fadhillah	
11	Effective Teaching in Mongolia: Policies, Practices and Challenges	245
	Amarjargal Adiyasuren and Ulziisaikhan Galindev	
12	An Assessment of the Learning Environment and Teacher Interpersonal Behaviour at the Teacher Education Level	257
	Adit Gupta and Priya Sharma	
13	Teaching Effectiveness in Spain: Towards an Evidence-Based Approach for Informing Policymakers	283
	Carmen-María Fernández-García, Mercedes Inda-Caro, and María-Paulina Viñuela-Hernández	
14	An Explanation of the ICALT Instrument's Measurement of Teaching Quality in Relation to Teacher Education and Policy in South Korea	299
	Seyeoung Chun, Okhwa Lee, and Deuk-Joon Kim	
15	Classroom Learning Environments and Assessment Practices in Science Classrooms in Western Australia	317
	Rekha B. Koul	
16	Teacher Effectiveness in Multiple Lenses: Secondary Analysis Lessons in the Measures of Effective Teaching Project	339
	James Ko, Zhijun Chen, Jieyan Celia Lei, and Ridwan Maulana	
Part III Effective Teaching: Comparison Across Countries		
17	Secondary Education Teachers' Effective Teaching Behaviour Across Five Countries: Does it Change Over Time?	361
	Ridwan Maulana, Amanda Maraschin Bruscato, Michelle Helms-Lorenz, Yulia Irnidayanti, Thelma de Jager, Ulziisaikhan Galindev, Amarjargal Adiyasuren, Abid Shahzad, Nurul Fadhillah, Seyeoung Chun, Okhwa Lee, Thys Coetzee, and Peter Moorer	

18 Teacher-Student Interactions: Theory, Measurement, and Evidence for Universal Properties That Support Students’ Learning Across Countries and Cultures 399
 Tara Hofkens, Robert C. Pianta, and Bridget Hamre

19 Affective Student–Teacher Relationships and Students’ Engagement: A Cross–Cultural Comparison of China and The Netherlands 423
 Debora Roorda, Mengdi Chen, and Marjolein Zee

20 The Mediated Relationship Between Secondary School Student Perceptions of Teaching Behaviour and Self-Reported Academic Engagement Across Six Countries 439
 Ridwan Maulana, Rikkert van der Lans, Michelle Helms-Lorenz, Sibel Telli, Yulia Irnidayanti, Nurul Fadhillah, Carmen-Maria Fernandez-Garcia, Mercedes Inda-Caro, Seyeoung Chun, Okhwa Lee, Thelma de Jager, and Thys Coetzee

21 Impact of Play-Based Pedagogies in Selected Asian Contexts: What Do We Know and How to Move Forward? 473
 Alfredo Bautista, Jimmy Yu, Kerry Lee, and Jin Sun

22 Effective Interpersonal Relationships: On the Association Between Teacher Agency and Communion with Student Outcomes 489
 Perry den Brok, Jan van Tartwijk, and Tim Mainhard

23 Exploring How Teachers’ Personal Characteristics, Teaching Behaviors and Contextual Factors Are Related to Differentiated Instruction in the Classroom: A Cross-National Perspective 509
 Annemieke Smale-Jacobse, Peter Moorer, Ridwan Maulana, Michelle Helms-Lorenz, Carmen-María Fernández-García, Mercedes Inda-Caro, Seyeoung Chun, Abid Shahzad, Okhwa Lee, Amarjargal Adiyasuren, Yulia Irnidayanti, Ulziisaikhan Galindev, and Nurul Fadhillah

Part IV Effective Teaching and Its Correlates

24 Teachers’ Intrinsic Orientation, Self-Efficacy, Background Characteristics, and Effective Teaching: A Multilevel Moderated Mediation Modeling 543
 Xiangyuan Feng, Michelle Helms-Lorenz, and Ridwan Maulana

25 The Effects of a Short Self-Access Online Training for Practicum Preparation on the Depths of Reflection of Preservice Teachers 575
 Ye Wang, James Ko, and Peng Wang

26	Effective and Inspiring Teaching in STEM Classrooms: Evidence from Classroom Observations with Instrument Comparisons	595
	James Ko	
27	Fostering Pupils' Deep Learning and Motivation in the Norwegian Context: A Study of Pupils' Perceptions of Mathematics Instruction and the Link to Their Learning Outcomes	619
	Inger Marie Dalehefte and Esther Tamara Canrinus	
28	The Influence of Science Teachers' Beliefs and Practices on Students' Learning Spaces and Processes: Insights from Singapore	635
	Yuen Sze Michelle Tan and Imelda Santos Caleon	
29	The Illusion of Perspective: Examining the Dynamic Between Teachers' Perceived and Observed Effective Teaching Behaviour . . .	653
	Benjamin Looker, Alison Kington, Kimberley Hibbert-Mayne, Karen Blackmore, and Scott Buckler	
Part V Effective Teaching in Complex Environments: Differentiation and Adaptive Teaching		
30	Differentiated Instruction as an Approach to Establish Effective Teaching in Inclusive Classrooms	677
	Esther Gheysens, Júlia Griful-Freixenet, and Katrien Struyven	
31	Evaluating Effective Differentiated Instruction in Multicultural South African Secondary Schools	691
	Thelma de Jager	
32	Dealing with the Complexity of Adaptive Teaching through Collaborative Teacher Professional Development	707
	Tijmen M. Schipper, Sui Lin Goei, and Siebrich de Vries	
33	Adapting Teaching to Students' Needs: What Does It Require from Teachers?	723
	Marieke van Geel, Trynke Keuning, Kyra Meutstege, Jitske de Vries, Adrie Visscher, Christel Wolterinck, Kim Schildkamp, and Cindy Poortman	
34	The Teacher's Turn: Teachers' Perceptions of Observed Patterns of Classroom Interaction	737
	Nienke Smit, Marijn van Dijk, Kees de Bot, and Wander Lowie	
35	How Do Dutch Teachers Implement Differentiation in Primary Mathematics Education?	757
	Emilie J. Prast and Marian Hickendorff	

**36 Differentiation and Students with Special Educational Needs:
Teachers' Intentions and Classroom Interactions 775**
Elisa Kupers, Anke de Boer, Judith Loopers, Alianne Bakker,
and Alexander Minnaert

Epilogue 793

Concluding Thoughts 797

About the Editors

Ridwan Maulana is an associate professor at the Department of Teacher Education, University of Groningen, the Netherlands. His major research interests include teaching and teacher education, factors influencing effective teaching, methods associated with the measurement of teaching, longitudinal research, cross-country comparisons, effects of teaching behaviour on students' motivation and engagement, and teacher professional development. He has been involved in various teacher professional development projects including the Dutch induction programme and school–university-based partnership. He is currently a project leader of an international project on teaching quality involving countries from Europe, Asia, Africa, Australia, and America. He is a European Editor of *Learning Environments Research* journal, a SIG leader of Learning Environments of American Educational Research Association, and chair of the Ethics Commission of the Teacher Education.

Michelle Helms-Lorenz is an Associate Professor at the Department of Teacher Education, University of Groningen, The Netherlands. Her research interest covers the cultural specificity versus universality (of behaviour and psychological processes). This interest was fed by the cultural diversity in South Africa, where she was born and raised. Michelle's second passion is education, the bumpy road toward development. Her research interests include teaching skills and well-being of beginning and pre-service teachers and effective interventions to promote their professional growth and retention.

Robert M. Klassen is Professor and Chair at the University of York in the UK. His research connects the areas of motivation, technology, and the teaching workforce. He is a Chartered Psychologist in the UK and a Fellow of the Academy of Social Sciences and the American Psychological Association

Chapter 1

Prologue



Ridwan Maulana , Michelle Helms-Lorenz , and Robert M. Klassen 

There is a growing desire to improve the quality and the equity of education around the world. Educational improvement requires understanding that the chief actors in the education system – teachers and students – and the educational context in which they operate, are indispensable in this pursuit. This book contributes to understanding educational systems and personal factors that influence teaching behaviour and student learning and engagement. Particularly, the book focuses on the work of teachers – in terms of effective teaching – as key players in education. Effective teaching refers to classroom processes or instructional practices related to student learning (Wagner et al., 2013). This broad definition encompasses various terms used in the literature on teaching to refer to similar constructs and ideas.¹ It is therefore important to note that the scope of this book represents various strands of research on teaching.

Although research on effective teaching has a rich history of over half a century, the knowledge base is still growing. Research on effective teaching has consistently revealed that in general, teachers' work is a significant factor for student learning and outcomes (Kyriakides et al., 2009). However, understanding the specific

¹ Other scholars use various terms such as *quality of teaching* (e.g., Hattie, 2009), *teaching quality* (e.g., Fauth et al., 2014), *teaching effectiveness* (e.g., Seidel & Shavelson, 2007), *classroom quality* (e.g., Hamre et al., 2014), *classroom management* (e.g., Arens et al., 2015), *classroom environment* (e.g., FraserDay et al., 2015), *classroom learning environment* (e.g., Fraser & Goh, 2003), *instructional quality* (e.g., Rjosk et al., 2014), *instructional style* (e.g., Jang et al., 2010), *teaching styles* (e.g., Wentzel, 2002), and *interpersonal teacher behaviour* (den Brok et al., 2004).

R. Maulana (✉) · M. Helms-Lorenz
Department of Teacher Education, University of Groningen, Groningen, The Netherlands
e-mail: r.maulana@rug.nl; m.helms-lorenz@rug.nl

R. M. Klassen
Department of Education, University of York, York, UK
e-mail: robert.klassen@york.ac.uk

conditions, specific interactions of teachers with specific students, and the underlying mechanisms that enhance learner engagement remain to be explored in more depth, as they require massive and perpetual endeavors to align with the dynamic nature of education in different settings. Studies on effective teaching have been dominated by developed, mostly Western, contexts (e.g., Australia, North America, The UK, and Europe). Extending the knowledge base beyond national borders by studying and sharing insights of education between more and less developed parts of the world, can foster reciprocal and global educational improvement.

This book aims to bring together theoretical, empirical, methodological, and practical insights from diverse countries and educational contexts on effective teaching. It particularly focuses on discussing issues pertaining to effective teaching behaviour including framing and conceptualizations, characteristics, measurements, antecedents, correlates, and importance to teacher and student outcomes from national perspectives. The book draws upon the rich cultures and diverse contexts around the globe including Asia, Australia, Africa, America, and Europe, in order to improve understanding of effective teaching from a wide spectrum of educational systems.

This book is not intended to supersede the existing excellent books in the field (e.g., Darling-Hammond et al., 2017; Hall et al., 2020; Kyriakides et al., 2018; Scherens, 2016). Rather, it aims to complement and extend the body of knowledge on teaching. This may be the first book documenting a wide variety of topics and rich contents related to effective teaching from such highly diverse international contexts. Particularly, the book presents research that is presently absent in the current literature. First, it integrates research on effective teaching from various frameworks, operationalisations, and professional development perspectives. Second, it presents contributions from various countries/cultures across five continents. Third, it includes a number of observation and survey studies on effective teaching across countries using the same instruments in the same classrooms (over time). Fourth, it represents various educational systems that vary in quality based on popular international testing studies. Fifth, it provides discussion about effective teaching from the perspectives of authors *in situ*, highlighting the scientific and practical implications for the specific as well as potential global contexts. Sixth, it includes various levels of education ranging from primary to tertiary education. Finally, the book also dedicates a section on differentiation and adaptive teaching that is currently gaining more popularity in education. The book is structured in five sections that each serve a different purpose.

Part I presents conceptualizations and measurements of effective teaching. Part II provides insights into effective teaching from various international contexts. Part III presents studies on effective teaching from various cultural contexts taking the comparative perspective. Part IV documents studies on effective teaching and its correlates. Part V compiles a number of studies on a contemporary issue in effective teaching: differentiation and adaptive teaching. This book closes with an Epilogue chapter drawing together insights and ideas discussed from Part I to Part V, taking into account commonalities and differences across the sections and chapters. Finally, this book closes with a Concluding chapter by the editors that provides

reflections and future directions for studies on effective teaching from international perspectives, and suggests potential recommendations for research, policy, and practice. The book can serve as a contemporary reference on effective teaching, with diverse content and research approaches that will be highly relevant in various scientific and educational programs across the world.

References

- Arens, A. K., Morin, A. J. S., & Watermann, R. (2015). Relations between classroom disciplinary problems and student motivation: Achievement as a potential mediator? *Learning & Instruction, 39*, 184–193. <https://doi.org/10.1016/j.learninstruc.2015.07.001>
- Darling-Hammond, L., Burns, D., Campbel, C., Goodwin, A. L., Hammerness, K., Low, E.-L., McIntyre, A., Sato, M., & Zeichner, K. (2017). *Empowered educators: How high-performing systems shape teaching quality around the world*. Jossey-Bass.
- Day, S. L., Connor, C. M., & McClelland, M. M. (2015). Children’s behavioral regulation and literacy: The impact of the first grade classroom environment. *Journal of School Psychology, 53*, 409–428. <https://doi.org/10.1016/j.jsp.2015.07.004>
- Den Brok, P., Brekelmans, M., & Wubbels, T. (2004). Interpersonal teacher behaviour and student outcomes. *School Effectiveness and School Improvement, 15*, 407–442. <https://doi.org/10.1080/09243450512331383262>
- Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning & Instruction, 29*, 1–9. <https://doi.org/10.1016/j.learninstruc.2013.07.001>
- Fraser, B. J., & Goh, S. C. (2003). Classroom learning environments. In J. P. Keeves et al. (Eds.), *International handbook of educational research in the Asia-Pacific region. Springer international handbooks of education, vol 11*. Springer. https://doi.org/10.1007/978-94-017-3368-7_32
- Hall, J., Lindorff, A., & Sammon, P. (2020). *International perspectives in in educational effectiveness research*. Springer.
- Hamre, B. K., Hatfield, B. E., Pianta, R. C., & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher-child interactions: Associations with preschool children’s development. *Child Development, 85*, 1257–1274. <https://doi.org/10.1111/cdev.12184>
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *Journal of Educational Psychology, 102*, 588–600. <https://doi.org/10.1037/a0019682>
- Kyriakides, L., Creemers, B. P. M., & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education, 25*(1), 12–23. <https://doi.org/10.1016/j.tate.2008.06.001>
- Kyriakides, L., Creemers, B., & Charalambous, E. (2018). *Equity and quality dimensions in educational effectiveness*. Springer.
- Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E., & Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality. *Learning and Instruction, 32*, 63–72. <https://doi.org/10.1016/j.learninstruc.2014.01.007>
- Scherens, J. (2016). *Educational effectiveness and ineffectiveness*. Springer.
- Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77*, 454–499. <https://doi.org/10.3102/0034654307310317>

- Wagner, W., Göllner, R., Helmke, A., Trautwein, U., & Lüdtke, O. (2013). Construct validity of student perceptions of instructional quality is high, but not perfect: Dimensionality and generalizability of domain-independent assessments. *Learning and Instruction*, 28, 1–11. <https://doi.org/10.1016/j.learninstruc.2013.03.003>
- Wentzel, K. R. (2002). Are effective teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development*, 73, 287–301. <https://doi.org/10.1111/1467-8624.00406>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Part I

Conceptualization and Measurement of Effective Teaching

Part I Overview

Part I of this book consists of seven chapters. These chapters represent a range of perspectives and provide a general background for studies on effective teaching situated in local and international contexts.

Chapter 2 starts with presenting a theory-driven and evidence-based approach for teaching – the dynamic model of educational effectiveness – and links educational effectiveness research with research on teaching improvement. The authors discuss the main elements of the dynamic model focusing on the classroom level factors and their measurement dimensions. Chapter 3 continues with discussing current conceptualizations, theories, measurements, and instruments of effective teaching, bringing together popular research strands including educational and teacher effectiveness, learning environments, and motivational theories. The chapter also presents important issues on effective teaching including contexts, antecedents, informants, and its dynamic characteristics. Chapter 4 presents a study about newly recruited teachers' performance, in terms of grade point average (GPA), for entry to the profession in Sweden over the last two decades. The study highlights a decrease in GPA for newly recruited teachers over time, and notes between-teacher variation depending on the certification status.

Chapter 5 presents a study from Canada on the use of a learning environment instrument called the Place-based and Constructivist Learning Environment Survey (PLACES) and links it to the development of students' citizenship values. The study sheds light on how paying close attention to the learning environment created within environmental education programming can contribute to long-term outcomes of active citizenship. Chapter 6 provides insights into measuring teacher effectiveness through student perceptions, discusses risks and opportunities of using student perceptions and the effective use of student feedback data for the development of teaching and teachers. Chapter 7 discusses the use of two observation instrument – ICALT and TEACH – for measuring effective teaching in under-advantaged province in China. The study concludes that these instruments cannot provide detailed accounts

of classroom processes, and argues that systematic qualitative analysis is indispensable to understand teacher evaluations based on observation instruments. Chapter 8 reports findings from South Korea on the use of an observation instrument – ICALT – for serving two purposes: the detection of teachers ‘current development’, and the identification of their zone of proximal development. The authors conclude that the observation instrument offers the possibility to coach teachers and guide them in practices that they are not yet implementing.

Chapter 2

Using Educational Effectiveness Research for Promoting Quality of Teaching: The Dynamic Approach to Teacher and School Improvement



Leonidas Kyriakides and Anastasia Panayiotou

Abstract The chapter discusses the need of using a theory-driven and evidence-based approach for teaching improvement purposes and argues that the dynamic model of educational effectiveness may be used for establishing links between educational effectiveness research and research on teaching improvement. In the first part of the chapter the main elements of the dynamic model are presented with an emphasis at the factors operating at classroom level and their measurement dimensions. The first part also provides an overview of national and international studies conducted to test the validity of the dynamic model at classroom level. These empirical studies have provided support for the importance of factors included in the dynamic model (such as application, modelling, student assessment etc.), with regard to their effects on student learning outcomes. Empirical studies have also revealed relationships among factors operating at the classroom level, which help us define stages of effective teaching. Therefore, in the second part of the chapter, we discuss ways of using the dynamic model for teaching improvement purposes. In this context, the rationale and main steps of the dynamic approach (DA) to teaching improvement are presented. In the final section, we provide a critical review of studies investigating the impact of the DA on improving teaching skills and promoting student learning outcomes and draw implications for research, policy, and practice.

Keywords Educational effectiveness research · Quality of teaching · Teacher professional development · Stages of effective teaching · Quality and equity in education

L. Kyriakides (✉) · A. Panayiotou
Department of Education, University of Cyprus, Nicosia, Cyprus
e-mail: kyriakid@ucy.ac.cy; anas.panayiotou@gmail.com

© The Author(s) 2023
R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_2

1 Introduction

Quality of teaching comprises a topic of interest for most educational systems around the world and actions for maximizing the effect of the teaching and learning processes on student learning outcomes are frequently undertaken by investing a significant amount of resources. However, many of the efforts made to improve quality of education may be considered fragmented, superficial and lacking theoretical and empirical support (Scheerens, 2013, 2016). Teacher training and professional development, which are considered essential mechanisms for improving quality of teaching through the development of teachers' teaching practices, is not always based on the existing knowledge-base. Teachers may thus be involved in professional development, the content of which was not found to be associated to student learning or their own individual needs for development (Creemers et al., 2013). Developing effective professional development programmes that can promote change in classroom practices (Darling-Hammond, 2000) is needed, so as to improve quality of teaching and, consequently, student learning outcomes. Teachers' improvement efforts should be based on a solid theoretical framework that has received empirical validation for its main assumptions and that may guide teachers' improvement efforts. Research within the field of Educational Effectiveness Research (EER) should, thus, be considered for designing professional development programmes that may lead to improvements in teaching practices (Kyriakides et al., 2020b). Towards that end, the Dynamic Approach to teaching improvement (DA) was developed and makes use of the dynamic model (Creemers & Kyriakides, 2008) which addresses the complexity of educational effectiveness, and at the same time, its representation of factors and measurement dimensions provide opportunities to design teaching improvement programmes which are flexible and differentiated to meet the needs of individual teachers situated at different stages (Creemers et al., 2013). More information on the DA may be found in Sect. 4. In this chapter, we acknowledge that variation exists in teacher effectiveness which should be taken into consideration when offering teacher professional development programmes (Antoniou, 2013; Muijs et al., 2014). The dynamic model, supports that the factors included at the teacher level can be classified into different stages of effective teaching, structured in a developmental order beginning from simpler teaching behaviour to more complex teaching skills (i.e., differentiation of teaching). In the next section, the rationale and main elements of the dynamic model are described.

2 The Dynamic Model of Educational Effectiveness

In this section the main elements and rationale upon which the dynamic model has been developed, are presented. The factors included at classroom level are analyzed and their main features are explained. Even though the dynamic model is multilevel in nature, in this chapter we only focus on the classroom level and present the

teaching factors as these have been systematically shown to have a greater effect on student learning than factors located at the upper levels (i.e., school and system). Despite the fact, that factors located at the upper levels also have effects on student outcomes, these are smaller and mostly indirect (Kyriakides et al., 2018b). Since, therefore, it would not be possible to equally address in this paper the factors of different levels, we place focus on the factors located at the classroom level. For more information on the factors included in the dynamic model at the upper and lower levels see Creemers and Kyriakides (2008).

2.1 Main Elements and Rationale

The dynamic model of educational effectiveness (Creemers & Kyriakides, 2008) depicts the outcome of a systematic attempt to develop a framework of effectiveness that is able to encompass the dynamic nature of education and that is comprehensive enough to be able to be used by stakeholders in education, in order to improve the outcomes of educational efforts. Namely, the main aim of its development was to establish links between EER and school improvement. The dynamic model was developed by considering the limitations of the integrated models of educational effectiveness and incorporated the findings of studies conducted regarding the factors that have an influence on student outcomes (Creemers & Kyriakides, 2008). It was developed based on the main principles of the Creemers' Comprehensive model (Creemers, 1994), however providing clearer definitions of the factors included at the different levels, as well as a more elaborated description of their measurement. In addition, the dynamic model takes into account the "new goals of education", which more broadly define the expected outcomes of schooling and are not restricted solely to the acquisition of basic skills. This means that apart from its reference to the cognitive outcomes of schooling, it also refers to other outcomes, such as affective, psychomotor, and new learning outcomes (e.g., metacognition). This portrays the need to view education in a more holistic manner and comprises ways of building upon previous theories of educational effectiveness. However, the dynamic model is based on the notion that a model should not only be parsimonious but should also be able to describe the complex nature of educational effectiveness. This implies that the model is based on a specific theory, but at the same time some of the factors included in the major constructs of the model are expected to relate to one another within and/or between levels. Therefore, the dynamic model is also multi-level in nature and refers to factors operating at the four levels shown in Fig. 2.1 (i.e., student, classroom, school, and system). However, special emphasis is placed at the classroom level and the roles of the two main actors (i.e., teacher and student) are analyzed.

The dynamic model also suggests that factors at the school and system level have both direct and indirect effects on student achievement since they are able to influence not only student achievement but also teaching and learning. In addition, the model assumes that there is a need to carefully examine the relationships between

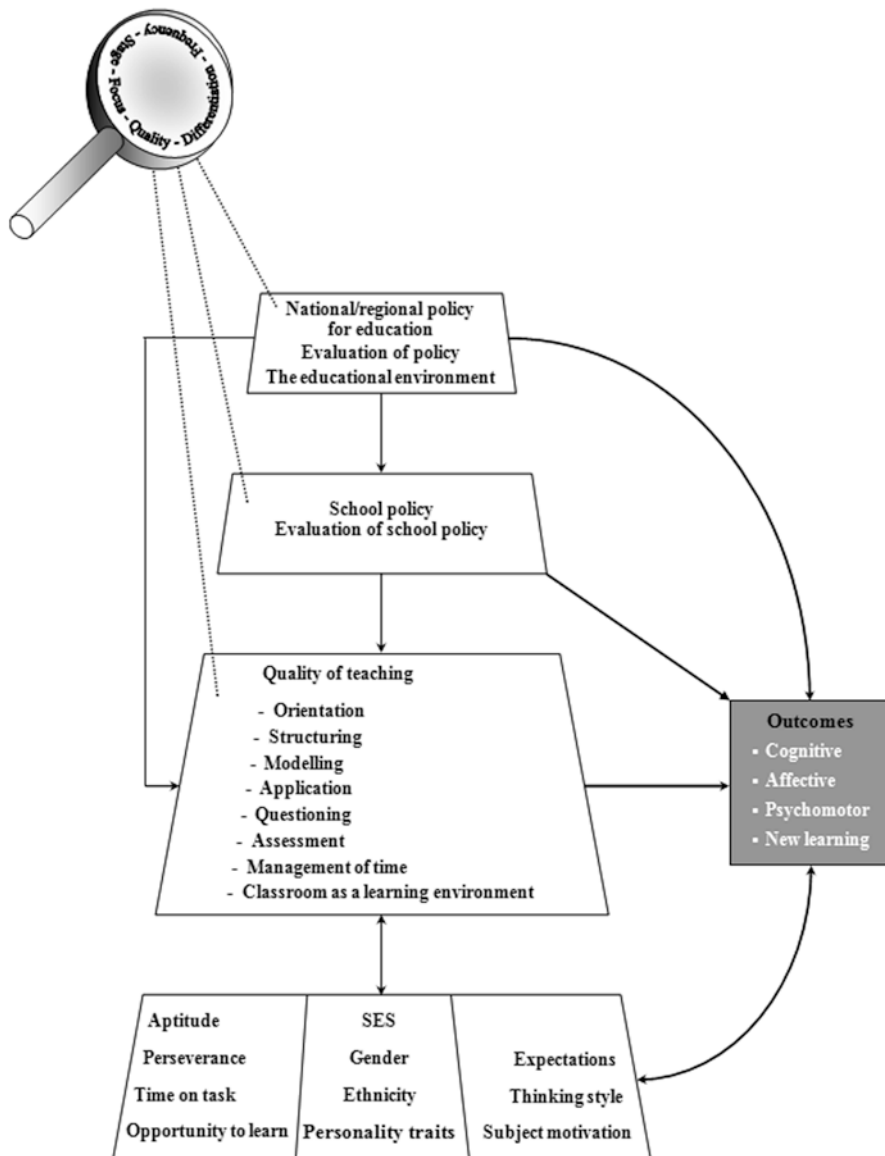


Fig. 2.1 The dynamic model of educational effectiveness

the various effectiveness factors which operate both at the same and different levels. Such relations were also demonstrated through earlier models such as Walberg's theory of educational productivity (Walberg, 1984) who indicated that aptitude, instruction and the psychological environment influence one another and are also influenced by feedback on the amount of learning that occurs. Such an approach to modelling educational effectiveness may reveal groupings of factors that make

teachers and schools more or less effective. Therefore, strategies for improving effectiveness which are comprehensive in nature may emerge. It should be noted here, that the dynamic model was designed in such way that can also be used for promoting improvement in education and not exclusively for research and theory development (Kyriakides et al., 2020b; Savage, 2012). In particular, the dynamic model aims to address another criticism made in the earlier theories of EER, regarding their practical use and the possibility of using their basic principles for policy development. The practical use of the dynamic model for improvement purposes, both at the classroom and school level, has been demonstrated through several experimental studies (for a review of these studies see Kyriakides et al., 2020b).

Finally, the dynamic model assumes that each factor can be defined and measured by using five dimensions: frequency, focus, stage, quality, and differentiation. This can be considered as one of the main differences of the dynamic model from all the existing theoretical models in EER, since other frameworks such as the Three Basic Dimensions of Teaching Quality (TBD) (Praetorius et al., 2018) and the International Comparative Analysis of Learning and Teaching” (ICALT) (Van de Grift, 2007), do not take into account the different dimensions with which factors may be measured. Therefore, the dynamic model attempts to show that effectiveness factors are multidimensional constructs and can be measured in relation to specific dimensions. The importance of taking each dimension of the teaching effectiveness factors into account is illustrated below.

- *Frequency* is a quantitative means of measuring the functioning of each factor. However, the other four dimensions which refer to the qualitative characteristics of the functioning of the factors reveal that effectiveness is more complicated than assumed by previous theoretical models and studies. Frequency, is probably the easiest way to measure the effect of a factor on student achievement, and, consequently, most effectiveness studies used this dimension to define effectiveness factors. For example, the frequency dimension of structuring is measured by taking into account the number of structuring tasks that take place in a typical lesson, as well as how long each structuring task takes place.
- *Focus* can be defined by taking into account two different facets. The first one refers to the specificity of the activities associated with the functioning of a factor, namely whether they can be considered as specific in terms of solid activities or policies; or more general, in terms of not providing adequate details to the different stakeholders on the application processes of an activity. The second aspect refers to the purpose for which an activity takes place by looking whether an action aims at achieving one or several purposes. The dynamic model argues that there should be a balance in the specificity of the teaching tasks and this assumption is in line with the synergy theory (see Kyriakides et al., 2020b). For example, with regard to the factor of structuring this may refer to the individual lesson or a series of lessons.
- *Stage* is related to the time at which tasks associated with a factor take place. It is assumed that the application of a factor in only one point in time may not constitute an effective way of dealing with the factor in terms of increasing the

positive effects resulting from its implementation. For example, structuring tasks are expected to take place not only at the beginning or end of a lesson, or unit of lessons, but at different time points so that the students are given the opportunity to develop links among the different parts of a lesson/series of lessons. Thus, the factors need to take place over a long period of time to ensure that they have a continuous direct or indirect effect on student learning.

- *Quality* refers to the properties of the specific factor itself, as they are discussed in the literature. For instance, in regard to the assessment factor, as it is stated through literature, formative assessment is expected to be more beneficial to students than summative and facilitate both learning and teaching (Black & Wiliam, 2009; Hattie & Timperley, 2007; Wiliam et al., 2004).
- *Differentiation* refers to the extent to which activities associated with a factor are applied without any digression for all the subjects involved with it (e.g., all the students, teachers, schools) irrespective of their needs and/or abilities. It is expected that adaptation to the specific needs of each subject or group of subjects will increase the successful implementation of a factor and will ultimately maximize its effect on student learning outcomes also addressing issues of equity (Kyriakides et al., 2018a). Taking in mind that students learn best when their teachers become accustomed to the differences in their readiness levels, interests and learning needs and make an effort to adjust their teaching in order to satisfy them (Tomlinson, 2005), the need for examining the functioning of the different factors in terms of differentiation is amplified. For example, teachers may ask students both process and product questions of different difficulty level, so as to give all students the opportunity to be engaged in a lesson.

In this section, the main assumptions and rationale upon which the dynamic model was developed were discussed. In the next section, a brief description of the factors included at classroom level is provided and their main characteristics are explained.

2.2 Teaching Factors: An Integrated Approach to Effective Teaching

Based on the main findings of teacher effectiveness research (e.g., Brophy & Good, 1986; Fraser et al., 1987; Muijs & Reynolds, 2001; Opdenakker & Van Damme, 2000; Rosenshine & Stevens, 1986), the dynamic model refers to factors which describe teachers' instructional role and are associated with student learning outcomes. These factors refer to observable instructional behaviour of teachers in the classroom rather than to factors that may explain such behaviour (e.g., teacher beliefs and knowledge and interpersonal competences). The eight factors included in the model are: orientation, structuring, questioning, teaching modelling, application, management of time, teacher role in making classroom a learning environment, and classroom assessment. These eight factors do not refer only to one approach of teaching, such as structured or direct teaching (Joyce et al., 2000) or to

approaches associated with constructivism (Schoenfeld, 1998). An integrated approach in defining quality of teaching is adopted. Specifically, the dynamic model does not refer only to skills associated with direct teaching and mastery learning such as structuring and questioning, but also to orientation and teaching modelling which are in line with theories of teaching associated with constructivism (Brekelmans et al., 2000). Moreover, the collaboration technique is included under the overarching factor of teacher contribution to the establishment of the classroom learning environment. Studies investigating differential teacher effectiveness have revealed that the previously listed eight factors may have a stronger impact on the learning of specific groups of students but can be treated as generic in nature as research has highlighted a link with the achievement of each group of students (Campbell et al., 2004). A short description of each factor follows. Information on the instruments for measuring these factors may, also, be found in Creemers and Kyriakides (2012).

- A) Orientation: This factor draws on theories in the field of motivation and refers to teacher behaviour in providing the students with opportunities to identify the reason(s) for which an activity or lesson or series of lessons occur and/or actively involving students to the identification of the reason(s) for which a lesson includes a specific task. Through this process it is expected that the activities that take place during lessons, become meaningful to students and consequently increase their motivation for participating actively in the classroom (e.g., De Corte, 2000; Paris & Paris, 2001). This factor may thus have an impact on increasing student motivation and through that, on increasing student learning outcomes.
- B) Structuring: Student learning is positively influenced when teachers actively present materials and structure them by: (a) beginning with overviews and/or review of objectives; (b) outlining the content to be covered and signaling transitions between lesson parts; (c) calling attention to main ideas; and (d) reviewing main ideas at the end (Rosenshine & Stevens, 1986). Structuring activities aim at assisting students develop links between the different parts of lessons, instead of dealing with them in an isolated way (Kyriakides et al., 2020b).
- C) Questioning: This factor is defined according to five elements. Firstly, effective teachers are expected to not only provide a large amount of product questions which require students to respond in a single way, but also focus on expecting students to elaborate on their answers and provide details on how they were able to reach their answer (i.e., by also posing process questions). Secondly, it is anticipated that teachers provide enough time to students before calling for their answers respective of each question's level of difficulty. Thirdly, the clarity of the questions posed is taken into consideration, so that no misconceptions or misinterpretations are caused. Fourthly, the question level of difficulty should reflect students' ability, avoiding too difficult questions that would inevitably cause complete failure to respond (Brophy & Good, 1986). Finally, it is outlined that an important aspect of this factor is the way teachers deal with student responses. Specifically, correct responses should be acknowledged so that all

students are aware of the correct answer at the end of the discussion. In case a student's answer is not fully correct then the teacher should acknowledge whatever part may be correct and assist the student in discovering the correct answer or provide an improved response, through the provision of clarification or helpful guidelines.

- D) **Teaching-modelling:** Although there is a long tradition in research on teaching higher-order thinking skills and problem solving, these teaching and learning activities have received unprecedented attention during the last two decades, due to the policy emphasis on the achievement of new goals of education. Thus, the teaching-modeling factor is associated with findings of effectiveness studies revealing that effective teachers are expected to help students use strategies and/or develop their own strategies that can help them solve different types of problems (Muijs et al., 2014). Consequently, students are expected to develop skills that help them organize their own learning (e.g., self-regulation and active learning). In defining this factor, the dynamic model also addresses the properties of teaching-modeling tasks, and the role that teachers are expected to play in order to help students devise problem-solving strategies. Teachers may either present students with a clear problem-solving strategy, or they may invite students to explain how they themselves would approach or resolve a particular problem and then use that information for promoting the idea of modeling. Recent research suggests that the latter approach may encourage students to not only use, but also develop their own problem-solving strategies (Aparicio & Moneo, 2005; Gijbels et al., 2006).
- E) **Application:** Providing students with practice and application opportunities can improve learning outcomes. Learning new information cannot be a constant process, since according to the Cognitive Load Theory the working memory can only process a limited amount of information at each given time (Kirschner, 2002). It is also argued that application tasks should not only constitute a repetition of the material that students were taught in classroom but should move a step forward adding more complex and mentally stimulating elements. Thus, application activities should provide the trigger for further knowledge, contributing to the linkage of the units taught in one lesson or series of lessons with the following. Effective teachers are expected to not only observe students engaging in application tasks, but also to actively contribute to their learning by supervising their progress and providing students with constructive feedback (Creemers et al., 2013).
- F) **The classroom as a learning environment:** This factor as described in the dynamic model consists of five components which were shown to be the most important aspects of the classroom climate through teacher effectiveness studies and meta-analyses: (a) teacher-student interaction, (b) student-student interaction, (c) students' treatment by the teacher, (d) competition between students, and (e) classroom disorder (Fraser & Goh, 2003). Classroom environment research has shown that the first two of these elements are important components of measuring classroom climate (see Cazden, 1986; Harjunen, 2012). However, the dynamic model is concerned with the immediate impact that

teacher initiatives have on establishing relevant interactions in the classroom, and it investigates the extent to which teachers are able to establish on-task behaviour through promotion of interactions. The other three elements refer to teachers' attempts to create an efficient and supportive environment for learning in the classroom (Walberg, 1986). These elements are measured by taking into account the teacher's behaviour in establishing rules, persuading students to respect and use the rules, and the teacher's ability to maintain them in order to create and sustain an effective learning environment in the classroom.

- G) Management of time: To address this factor the amount of time used per lesson for on-task behaviour is investigated. Teachers are expected to: (a) prioritize academic instruction and allocate available time to curriculum-related activities; and (b) maximize student engagement rates. Time management skills are not restricted solely to teachers' ability to avoid the loss of teaching time through minimizing external classroom disruptions, or through dealing effectively with organizational issues (e.g., moving between classes, organizing and distributing materials or giving instructions). Apart from the overall teaching time, management of time skills also include teacher actions that increase the learning time for each individual student (i.e., the on-task time).
- H) Assessment: Assessment is seen as an essential part of teaching (Stenmark, 1992). Especially formative assessment has been shown to be one of the most important factors associated with effectiveness at all levels, especially at the classroom level (Christoforidou et al., 2014). Effective teachers are therefore expected to: (a) Use appropriate techniques to collect data on student knowledge and skills; (b) analyze data in order to identify student needs; (c) report assessment results to students and parents; and (d) evaluate their own practices.

In this section, the factors included at the classroom level of the dynamic model have been briefly described, in the next section, a description of the main studies that have provided empirical support to the main assumptions of the model at the classroom level is provided.

3 Empirical Support Provided to the Main Assumptions of the Dynamic Model at the Classroom Level

Sixteen empirical studies have been conducted thus far to examine the main assumptions of the dynamic model at classroom level. These studies have been able to demonstrate that teaching factors in the dynamic model are associated with students' achievement gains. It is also important to note that different types of learning outcomes were used as criteria for measuring teacher effectiveness. Namely, the impact of teaching factors was demonstrated on promoting not only cognitive, but also affective (e.g., Kyriakides & Creemers, 2008) psychomotor (e.g., Kyriakides et al., 2018c) and meta-cognitive learning outcomes (e.g., Kyriakides et al., 2020a). Different subjects (i.e., language, mathematics, science, religious education, and

physical education) and different phases of education (i.e., pre-primary, primary, and secondary education), have also been considered in these studies. Therefore, these studies provided some empirical support for the assumption that teaching factors can be generic. However, it should be noted that only two studies examined the impact of the teaching factors on non-cognitive outcomes and only one on student metacognitive outcomes. What is, however, more important is that in some studies it was not possible to see the effects of some factors when only the frequency dimension was considered, but variation in student achievement was explained when the other four dimensions of these factors were taken into account (e.g., Kyriakides et al., 2020b). It is relevant to point out that one of these studies was conducted in Ghana whereby the observation instruments and the student questionnaire were used to collect data on the teaching factors of the dynamic model and measure the impact of teaching factors on mathematical achievement of primary students in Ghana (see Azigwe et al., 2016). In this study no effect of the teaching factors was identified through the student questionnaire which was able to collect data on all eight teaching factors but not on all measurement dimensions and therefore only the data collected through the observation instruments were used to measure the effect of the teaching factors on student achievement. This shows the need to also collecting observational data for the measurement of the factors. Similar results were also found in a study in the Maldives where data collected through the student questionnaire were able to detect the effect of only few factors on student learning outcomes whereas observation data were able to detect the effect of all factors on student learning outcomes (Musthafa, 2020).

Regarding the link between effectiveness factors and their impact on student achievement, Kyriakides et al. (2013) conducted a quantitative synthesis of 167 studies, which had been carried out between 1980 and 2010 and which had been designed to investigate the contribution of teacher classroom behaviours to student learning outcomes. For the purpose of this synthesis, all the selected studies included explicit and valid measures of student achievement in relation to cognitive, affective or psychomotor outcomes of schooling. Studies that used more global criteria for academic outcomes, such as dropout rates, grade retention and enrolment in universities, were also included. Given the focus of this meta-analysis, a study was included if it also had measures of specific teaching factors and provided information on the methods used to measure each factor. This meta-analysis not only revealed that factors included in the dynamic model were moderately associated with student achievement, but also that the type of outcomes had no significant effect on the functioning of the factors examined in the study. On the other hand, the type of study did have an effect since experimental studies were found to report higher effect sizes than longitudinal and cross-sectional studies. This meta-analysis, also, revealed that factors not included in the model were weakly associated with student learning, except for concept mapping and self-regulation. However, the effect of concept mapping was only investigated through three studies which were experimental in nature, hence the strong average size reported for concept mapping should not be dissociated from the nature of the studies considered with respect to this factor. With regard to self-regulation, this may be seen as closely associated to

other factors already included in the dynamic model. For example, the orientation factor included in the model attends to the extent to which the teacher provides information to orient students towards the importance of learning the new content. This factor of the dynamic model could be considered as a component of teachers' attempt to encourage self-regulation and help students understand the reasons for which they should be engaged in certain learning tasks. From a theoretical standpoint, then, such connections suggest that including self-regulation in the dynamic model might be a natural extension to the model. This is because this factor can help better capture the extent to which teaching not only gives students the opportunity to apply approaches presented in the lesson (i.e., application) or to develop certain strategies for dealing with particular problems (i.e., modelling), but it can also help students gradually become independent learners.

Finally, the findings of this meta-analysis provide some empirical support for the use of an integrated approach to defining effective teaching, especially since the factors found to have an effect on student outcomes, be they (meta) cognitive, affective or psychomotor, were not associated solely with either the direct and active teaching approach or the constructivist approach. For example, this meta-analysis showed that factors related to direct instruction (e.g., time management, structuring) or to constructivism (e.g., orientation, modelling) both contribute to student learning outcomes. This finding empirically supports the assumptions of the dynamic model, which, pursues an integrated approach and incorporates factors from different instructional perspectives at the teacher/classroom level (see Kyriakides, 2008).

Despite the abovementioned studies and meta-analysis, it should be noted that, no analyses have been done to examine whether the factors may be grouped into second order overarching factors, however, studies have supported the assumption that the teaching factors of the dynamic model and their dimensions are inter-related and revealed that they can be classified into stages of effective teaching, structured in a developmental order by using the Rasch model (see Kyriakides et al., 2020b).

In particular, the first study that revealed relationships among the teaching factors (Kyriakides et al., 2009) was conducted to identify the impact of the eight teaching factors and their dimensions on student achievement gains in different subjects (i.e., language, mathematics, and religious education) and on different types of learning outcome (i.e., cognitive and affective). This study tested the validity of the measurement dimension framework proposed by the dynamic model and made use of the Rasch model to identify the extent to which the five dimensions of the teaching factors could be reducible to a common unidimensional scale. By analyzing the data that emerged from the observation instruments used to measure the performance of the teacher sample in relation to the eight teaching factors and their dimensions, it was discovered that the data fitted the Rasch model, and a reliable hierarchical scale of teaching skills was established. Then, by using cluster analysis, it was found that the teaching skills could be grouped into five levels of difficulty that could be taken to stand for different types of teacher behaviour, moving from relatively easy to more difficult and spanning the five dimensions of the eight teaching factors included in the dynamic model.

The first three levels are mainly related to the direct and active teaching approach, moving from the basic requirements concerning quantitative characteristics of teaching routines to the more advanced requirements concerning the appropriate use of these skills as measured by the qualitative characteristics of these factors. These skills also gradually move from the use of teacher-centered approaches to the active involvement of students in teaching and learning. The last two levels are more demanding since teachers are expected to differentiate their instruction (level 4) and to demonstrate their ability to use the new teaching approach (level 5). Multilevel analysis of student achievement also showed that teachers situated at higher levels are more effective than those situated at the lower levels. This association is found with respect to achievement in all three different subjects and both cognitive and affective outcomes (see Kyriakides et al., 2009).

Similar results emerged from a study conducted in Canada which made use of student ratings to measure the skills of teachers in relation to each teaching factor and its dimensions (Kyriakides et al., 2013). In this case the stages which were identified also moved gradually from skills associated with direct teaching to more advanced skills involved in the constructivist approach and differentiation of teaching. This indicates that teachers may also move gradually from one type of teaching behaviour to a more complex one. An experimental study also investigated the impact of offering the teaching improvement programmes based on the dynamic approach for a longer period rather than just a single school year (Kyriakides et al., 2017). This study revealed that a stepwise progression of teachers' skills took place (over a period of three school years) and thus supported the generalizability of findings of the studies seeking to identify stages of effective teaching.

4 Establishing Links Between Theory and Practice: The Dynamic Approach to Teaching and School Improvement

The dynamic model has been developed taking into consideration that the theoretical base of educational effectiveness research should provide a basis for policy development and guide teaching and school improvement efforts. It is argued that in many cases, the relationship between science and practice in education and in educational effectiveness, specifically, has not been successful (Kyriakides et al., 2020b). However, considering research evidence when designing and implementing improvement programmes in education may lead to better student outcomes that reflect the efforts of practitioners towards improvement. Therefore, this chapter argues that the dynamic model may contribute to establishing a theory-driven and evidence-based approach to teacher professional development.

Regarding teacher professional development, different approaches are used, which in many cases, however, do not consider existing knowledge on effective teaching and the ways that teachers could better learn and implement educational practices that were found to be effective in promoting learning (Borko et al., 2010).

In this context, it is acknowledged that in the literature of teacher professional development, different views exist on the methodology, structure, and philosophical perspectives of different approaches to teacher training and professional development and the role of teachers in the developmental process (Day & Sachs, 2005). Towards that end, research on teacher training and professional development indicates two dominant approaches which may be seen not only as different, but also as rather opposing: the Competency-Based Approach (CBA) and the Holistic Approach (HA). On one hand, the CBA emphasizes skill acquisition through the setting of professional standards for teachers. Such professional standards have been developed on the assumption that it is possible to define what teachers should know and, most importantly, be able to do. This approach has been criticized for reinforcing teachers' practices in a reproductive way separating practice from content and restricting teachers' critical and creative thinking (Sprinthal et al., 1996). On the other hand, the HA which recognizes reflection as the way for teachers to develop effective practice has also been extensively criticized. Whereas reflection is identified as an important element in all aspects of learning (Ottesen, 2007); contradictory interpretations of what constitutes reflection (Cornford, 2002; Fendler, 2003) and how it translates into action (Cornford, 2002) can be identified. What is most important, however, is that none of these dominant approaches has provided enough evidence of their positive effect on teaching and learning. Taking the above mentioned into consideration, the Dynamic Approach (DA) to teacher professional development was proposed (Creemers et al., 2013) in an attempt to link EER with research on teacher professional development and address the limitations of the currently employed professional development approaches.

First, the DA assumes that teacher improvement efforts should aim at the development of teaching skills which relate to positive student learning outcomes. It is argued that teaching skills should not be addressed separately through teacher professional development without considering the professional needs of teachers (as proposed by the CBA) or very broadly (as implied by the HA) but rather, teacher training and professional development should address specific groupings of teaching factors in relation to student learning. Therefore, the DA draws on the two dominant approaches (i.e. the CBA and the HA) and aims to overcome their main weaknesses through considering the grouping of teaching factors included in the dynamic model of educational effectiveness (Creemers & Kyriakides, 2008). The main steps of the DA to teacher development are presented next.

4.1 The Main Steps of the DA

This section refers to the four main steps of the DA. The first step is concerned with the identification of the professional development needs of each teacher separately through empirical investigation. The DA assumes that an initial evaluation of teachers' teaching skills should be conducted prior to offering teacher training, to investigate the extent to which they possess certain teaching skills while identifying their needs and priorities for improvement (Creemers et al., 2013). The results of the

initial evaluation can help us classify teachers into developmental stages of teaching and generate suggestions for the content of training to be offered to different groups of teachers based on the stage at which they were found to be situated. The second step is concerned with the support that the advisory team (i.e. mentors) will provide to teachers in order to help them establish their own action plans. Specifically, the advisory team is expected to provide teachers of each group with supporting literature and research findings related to the teaching skills of their developmental stage. As a result, each teacher is in a position to develop his/her own action plan. The next step of the DA comprises the establishment of formative evaluation procedures. The formative evaluation procedures refer to the identification of the learning goals, intentions or outcomes and criteria for achieving them; the provision of timely and constructive feedback to enable teachers advance their learning; the active involvement of teachers in their own learning and, lastly, improvement in teaching skills. These procedures could be accomplished by the close collaboration of the advisory team and the participating teachers. The final step of the DA aims to identify the impact of the teacher professional development programme on the development of teachers' skills and its indirect effect on student learning. The results of summative evaluation assist in measuring the effectiveness of the DA and allow subsequent decisions to be made on how to further improve the programme and maximize its effect on educational quality. In the next section, experimental studies investigating the impact of this approach on improving teaching and promoting student learning outcomes are briefly presented.

4.2 Research on the Impact of the DA on Improving Teaching and Promoting Student Learning

Recent studies support the effectiveness of the DA in relation to the CBA and the HA. Particularly, a group randomisation study compared the effectiveness of the DA to the HA (Antoniou & Kyriakides, 2011). A total number of 130 teachers volunteered to participate in a teacher professional development programme. Their teaching skills and achievement of their students in mathematics ($n = 2356$) were measured at the beginning and at the end of the intervention. Teachers found to be at each developmental stage at the beginning of the intervention were randomly allocated evenly into two groups. The first group employed the DA and the second the HA. Teachers employing the DA managed to improve their teaching skills more than teachers employing the HA. The use of the DA also had a significant impact on student achievement gains in mathematics. In addition, all teachers of the study, participated in a follow-up measurement of their teaching skills, which took place 1 year after the end of the intervention. One year after the end of the intervention, the teaching skills of the participating teachers were evaluated using the same procedures as those used to measure their skills at the beginning and end of the intervention. The aim of this follow-up study was to investigate whether teachers had fallen back to their initial stage or whether they had continued to improve their teaching skills even after the intervention stimulus had ended. Analyses of data

provided evidence to compare the impact of the two approaches to TPD 1 year after the end of the intervention. Regarding the sustainability of the intervention, the follow-up measurement of teaching skills 1 year after the end of the interventions revealed no further improvement or decline in the teaching skills of either the DA or the HA group (Antoniou & Kyriakides, 2013). Taking into consideration, the improvement of teaching skills on the part of the DA group during the intervention, we argue that teachers can improve their teaching skills when they are exposed to appropriate interventions and participate in effective and systematic professional development programmes. Research findings also support the view that improvement is more apparent in those teachers who continue with informal education and participate systematically in effective professional development programmes (e.g., King & Kitchener, 1994). This is an important reminder that stage growth does not develop spontaneously but requires a stimulating and supportive environment. This project seems to reveal that such an environment can be established when teaching improvement projects based on the DA are offered to teachers. The second study compared the effectiveness of the DA to the CBA in improving teacher assessment skills and promoting student outcomes. Following the same approach as in the first study, teachers were invited to participate in a professional development programme and their skills in conducting assessment as well as the achievement of their students in mathematics ($n = 2358$) were measured at the beginning and at the end of the intervention. Teachers found to be at a certain stage at the beginning of the intervention were again randomly allocated evenly into two groups (see Christoforidou et al., 2014). The first group employed the DA and the second the CBA. The results of the study demonstrated that, for teachers at all stages, the DA was more effective in improving both assessment skills, as well as student outcomes in mathematics (see Creemers et al., 2013). Since experimental studies demonstrated that one-year interventions based on the DA have a positive impact on teacher effectiveness, a study took place by Kyriakides et al. (2017) aimed to examine the impact that a long-term programme based on the DA may have on quality of teaching. Therefore, a three-year school-based professional development programme was offered to 106 in-service primary education teachers in Cyprus coming from different public schools. Particularly, in-service primary school teachers were randomly allocated into two groups. The first group received a three-year programme based on the DA whereas the second acted as the control group. Pre- and post-measurement of teaching skills were performed each year. Results showed that, offering the DA for a longer period resulted in bigger effects on improving teaching skills but no change in the skills of the control group was observed. Namely, the effect sizes measuring the impact of offering the DA for 1 year (0.17), 2 years (0.30) or 3 years (0.39) reveal that the duration of a programme based on the DA plays an important role in improving teaching skills. During the first year of the implementation of the project a small effect of the DA on improving teaching skills was identified which is a similar result to those reported in previous studies investigating the impact of offering the DA for only 1 year. However, by offering the DA for a period of 3 years a bigger effect on improving teaching was identified which provides implication for the duration of teacher professional development.

5 Conclusion – Global Perspectives of Educational Effectiveness

EER has significantly evolved during the past decades both in terms of methodology, as well as, in terms of theory. The significance of teaching factors as the most important predictor of student learning outcomes, has also been systematically demonstrated (Muijs et al., 2014; Scheerens and Bosker, 1997). However, most studies have been conducted in developed- western countries, with a significantly smaller amount having been conducted in developing, and particularly SubSaharan African countries, which portray significant differences in contextual variables (Riddell, 2008). Research evidence suggest that teachers and schools may matter more in developing rather than in developed countries. Namely, a recent study conducted in Ghana (Azigwe et al., 2016), revealed that 55 per cent of the total variance in student achievement in mathematics was situated at the classroom level and only 45 per cent at the student level. This finding suggests that the classroom/teacher effect is much bigger in Ghana than in developed countries where studies conducted during the last four decades reveal that more than 60 per cent of variance is situated at the student level (see Creemers & Kyriakides, 2008; Scheerens & Bosker, 1997). Therefore, examining the differences in teacher effectiveness in different countries around the world and especially developing countries, is essential in terms of not only achieving quality of teaching in different educational settings, but also addressing issues of equity and equal opportunities in education and learning. In addition, using cross sectional data, Heyneman and Loxley (1982) found that SES was more important than school factors in determining children's academic performance in economically developed countries. Similar results are reported by Park (2008), who discussed how the association of the home literacy environment on reading achievement varies from country to country. Therefore, cross-national studies are needed to examine the effects of different factors in different educational settings. In addition, EER has frequently been criticized as being developed apart from teaching practice. Similarly, the results of teacher effectiveness research have not always provided a basis for teacher improvement efforts. Despite the improvements made to the field of EER during the last three decades, regarding research design, improvements in sampling techniques, and improvements in statistical techniques, the link between EER and professional development is still problematic. For this reason, we propose the establishment of strategies for teacher improvement which give emphasis on the evidence stemming from theory and research. Thus, the value of a theory-driven approach to teacher professional development is stressed. To that end, the DA was developed that considers the individual teacher professional development needs of teachers and is based on the assumption that teacher improvement efforts should aim at the development of teaching skills which were found to be related to improved student learning outcomes. Moreover, the DA aims to address the main weaknesses of the two dominant approaches (i.e., the CBA and the HA) to teacher professional development by considering the inter-relations between effectiveness factors when designing teacher training. Even though studies have shown the impact of the DA

on improving teaching skills and student learning outcomes, the sustainability of the results of the DA after the intervention need further investigation. One experimental study attempted to examine the one-year sustainability of the effects of the DA to teacher professional development (see Antoniou & Kyriakides, 2013) and revealed that, one year after the end of the interventions, no further improvement or decline in the teaching skills of the participating teachers took place. This may be partly explained by the fact that teaching experience alone without any form of teacher professional development does not contribute to the improvement of teaching skills (Çakir & Bichelmeyer, 2016; Huang & Moon, 2009). Taken that most recent studies on teacher professional development examine the short terms effects of providing teachers with professional development and even if positive effects are observed the sustainability of these effects is not determined (Derri et al., 2015), more research is needed to examine issues of sustainability of the effects of the DA.

Despite issues of sustainability, one should also examine the role of the Advisory and Research Team (A&R Team) that the DA assumes to have an important role towards the improvement of teaching skills. This team, consisting of researchers on teacher effectiveness and teacher professional development experts, is able to make available the appropriate knowledge base on improving the teaching skills that are set as improvement priorities for each teacher, as well as possessing technical expertise. The A&R Team is also expected to facilitate the process of formative assessment which is foreseen by the DA for monitoring the actions undertaken. Therefore, the degree to which the support of the A&R Team is needed for teacher improvement purposes, as well as the contribution of establishing formative assessment mechanisms, should also be examined. Finally, it should be acknowledged that studies examining the impact of the DA were only focused on determining its effect on improving student outcomes and have not dealt with issues of equity in education (Kyriakides et al., 2018a). Therefore, more studies are needed that search for the impact of DA on not only promoting student learning outcomes but also contributing to the reduction of the impact of background factors on student learning outcomes. These studies may help us identify how teacher professional development programmes can contribute in promoting both the quality and equity in education.

References

- Antoniou, P. (2013). A longitudinal study investigating relations between stages of effective teaching, teaching experience, and teacher professional development approaches. *Journal of Classroom Interaction*, 48(2), 25–40.
- Antoniou, P., & Kyriakides, L. (2011). The impact of a dynamic approach to professional development on teacher instruction and student learning: Results from an experimental study. *School Effectiveness and School Improvement*, 22(3), 291–311. <https://doi.org/10.1080/09243453.2011.577078>
- Antoniou, P., & Kyriakides, L. (2013). A dynamic integrated approach to teacher professional development: Impact and sustainability of the effects on improving teacher behavior and student outcomes. *Teaching and Teacher Education*, 29(1), 1–12. <https://doi.org/10.1016/j.tate.2012.08.001>

- Aparicio, J. J., & Moneo, M. R. (2005). Constructivism, the so-called semantic learning theories, and situated cognition versus the psychological learning theories. *Spanish Journal of Psychology*, 8(2), 180–198. <https://doi.org/10.1017/s1138741600005060>
- Azigwe, J. B., Kyriakides, L., Panayiotou, A., & Creemers, B. P. M. (2016). The impact of effective teaching characteristics in promoting student achievement in Ghana. *International Journal of Educational Development*, 51, 51–61. <https://doi.org/10.1016/j.ijedudev.2016.07.004>
- Black, P., & Wiliam, D. (2009). Developing a theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- Borko, H., Jacobs, J., & Koellner, K. (2010). Contemporary approaches to teacher professional development. *International Encyclopedia of Education*, 7(2), 548–556.
- Brekelmans, M., Slegers, P., & Fraser, B. J. (2000). Teaching for active learning. In P. R. J. Simons, J. L. van der Linden, & T. Duffy (Eds.), *New learning* (pp. 227–242). Kluwer.
- Brophy, J., & Good, T. L. (1986). Teacher behaviour and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328–375). Macmillan.
- Çakır, H., & Bichelmeyer, B. A. (2016). Effects of teacher professional characteristics on student achievement: An investigation in blended learning environment with standards-based curriculum. *Interactive Learning Environments*, 24(1), 20–32. <https://doi.org/10.1080/10494820.2013.817437>
- Campbell, R. J., Kyriakides, L., Muijs, R. D., & Robinson, W. (2004). *Assessing teacher effectiveness: A differentiated model*. Routledge Falmer.
- Cazden, C. B. (1986). Classroom discourse. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 432–463). Macmillan.
- Christoforidou, M., Kyriakides, L., Antoniou, P., & Creemers, B. P. M. (2014). Searching for stages of teacher skills in assessment. *Studies in Educational Evaluation*, 40, 1–11. <https://doi.org/10.1016/j.stueduc.2013.11.006>
- Cornford, I. R. (2002). Reflective teaching: Empirical research findings and some implications for teacher education. *Journal of Vocational Education & Training*, 54(2), 219–236. <https://doi.org/10.1080/13636820200200196>
- Creemers, B. P. M. (1994). *The effective classroom*. Cassell.
- Creemers, B. P. M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. Routledge.
- Creemers, B. P. M., & Kyriakides, L. (2012). *Improving quality in education: Dynamic approaches to school improvement*. Routledge.
- Creemers, B. P. M., Kyriakides, L., & Antoniou, P. (2013). Teacher professional development for improving quality in teaching. *Springer*. <https://doi.org/10.1007/978-94-007-5207-8>
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, 8(1). Retrieved from <http://epaa.asu.edu/epaa/v8n1/>
- Day, C., & Sachs, J. (2005). *International handbook on the continuing professional development of teachers*. McGraw-Hill Education.
- De Corte, E. (2000). Marrying theory building and the improvement of school practice: A permanent challenge for instructional psychology. *Learning and Instruction*, 10(3), 249–266. [https://doi.org/10.1016/s0959-4752\(99\)00029-8](https://doi.org/10.1016/s0959-4752(99)00029-8)
- Derri, V., Vasiliadou, O., & Kioumourtzoglou, E. (2015). The effects of a short-term professional development program on physical education teachers' behaviour and students' engagement in learning. *European Journal of Teacher Education*, 38(2), 234–262. <https://doi.org/10.1080/002619768.2014.947024>
- Fendler, L. (2003). Teacher reflection in a hall of mirrors: Historical influences and political reverberations. *Educational Researcher*, 32(3), 16–25. <https://doi.org/10.3102/0013189X032003016>
- Fraser, B. J., & Goh, S. C. (2003). Classroom learning environments. In *International handbook of educational research in the Asia-Pacific region* (pp. 463–475). Springer.

- Fraser, B. J., Walberg, H. J., Welch, W. W., & Hattie, J. A. (1987). Syntheses of educational productivity research. *International Journal of Educational Research*, 11(2), 145–252.
- Gijbels, D., Van de Watering, G., Dochy, F., & Van den Bossche, P. (2006). New learning environments and constructivism: The students' perspective. *Instructional Science*, 34(3), 213–226. <https://doi.org/10.1007/s11251-005-3347-z>
- Harjunen, E. (2012). Patterns of control over the teaching-studying-learning process and classroom as complex dynamic environments: A theoretical framework. *European Journal of Teacher Education*, 35(2), 139–161. <https://doi.org/10.1080/02619768.2011.643465>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Heyneman, S. P., & Loxley, W. A. (1982). Influences on academic achievement across high and low income countries: A re-analysis of IEA data. *Sociology of Education*, 55(1), 13–21.
- Huang, F. L., & Moon, T. R. (2009). Is experience the best teacher? A multilevel analysis of teacher characteristics and student achievement in low performing schools. *Educational Assessment, Evaluation and Accountability*, 21(3), 209–234. <https://link.springer.com/article/10.1007/s11092-009-9074-2>
- Joyce, B. R., Weil, M., & Calhoun, E. (2000). *Models of teaching*. Allyn & Bacon.
- King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. Jossey-Bass.
- Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning [Editorial]. *Learning and Instruction*, 12(1), 1–10. [https://doi.org/10.1016/S0959-4752\(01\)00014-7](https://doi.org/10.1016/S0959-4752(01)00014-7)
- Kyriakides, L. (2008). Testing the validity of the comprehensive model of educational effectiveness: A step towards the development of a dynamic model of effectiveness. *School Effectiveness and School Improvement*, 19(4), 429–446. <https://doi.org/10.1080/09243450802535208>
- Kyriakides, L., & Creemers, B. P. M. (2008). Using a multidimensional approach to measure the impact of classroom level factors upon student achievement: A study testing the validity of the dynamic model. *School Effectiveness and School Improvement*, 19(2), 183–205. <https://doi.org/10.1080/09243450802047873>
- Kyriakides, L., Creemers, B. P. M., & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education*, 25(1), 12–23. <https://doi.org/10.1016/j.tate.2008.06.001>
- Kyriakides, L., Archambault, I., & Janosz, M. (2013). Searching for stages of effective teaching: A study testing the validity of the dynamic model in Canada. *Journal of Classroom Interaction*, 48(2), 11–24. <https://www.jstor.org/stable/43858891>
- Kyriakides, L., Christoforidou, M., Panayiotou, A., & Creemers, B. P. M. (2017). The impact of a three-year teacher professional development course on quality of teaching: Strengths and limitations of the dynamic approach. *European Journal of Teacher Education*, 40(4), 465–486. <https://doi.org/10.1080/02619768.2017.1349093>
- Kyriakides, L., Creemers, B. P. M., & Charalambous, E. (2018a). Equity and quality dimensions in educational effectiveness. *Springer*. https://doi.org/10.1007/978-3-319-72066-1_1
- Kyriakides, L., Georgiou, M. P., Creemers, B. P. M., Panayiotou, A., & Reynolds, D. (2018b). The impact of national educational policies on student achievement: A European study. *School Effectiveness and School Improvement*, 29(2), 171–203. <https://doi.org/10.1080/09243453.2017.1398761>
- Kyriakides, E., Tsangaridou, N., Charalambous, C., & Kyriakides, L. (2018c). Integrating generic and content-specific teaching practices in exploring teaching quality in primary physical education. *European Physical Education Review*, 24(4), 418–448. <https://doi.org/10.1177/1356336x16685009>
- Kyriakides, L., Anthimou, M., & Panayiotou, A. (2020a). Searching for the impact of teacher behavior on promoting students' cognitive and metacognitive skills. *Studies in Educational Evaluation*, 64. <https://doi.org/10.1016/j.stueduc.2019.100810>

- Kyriakides, L., Creemers, B. P. M., Panayiotou, A., & Charalambous, E. (2020b). *Quality and equity in education: Revisiting theory and research on educational effectiveness and improvement*. Routledge.
- Muijs, D., & Reynolds, D. (2001). *Effective teaching: Evidence and practice*. Sage.
- Muijs, R. D., Kyriakides, L., van der Werf, G., Creemers, B. P. M., Timperley, H., & Earl, L. (2014). State of the art-teacher effectiveness and professional learning. *School Effectiveness and School Improvement*, 25(2), 231–256. <https://doi.org/10.1080/09243453.2014.885451>
- Musthafa, H. S. (2020). *A longitudinal study on the impact of instructional quality on student learning in primary schools of Maldives*. University of Cyprus. (Unpublished doctoral dissertation).
- Opendakker, M. C., & Van Damme, J. (2000). Effects of schools, teaching staff and classes on achievement and Well-being in secondary education: Similarities and differences between school outcomes. *School Effectiveness and School Improvement*, 11(2), 165–196. <https://www.tandfonline.com/doi/abs/10.1076/0924-3453%28200006%2911%3A2%3B1-Q%3BFT165>
- Ottesen, E. (2007). Reflection in teacher education. *Reflective Practice*, 8(1), 31–46. <https://doi.org/10.1080/14623940601138899>
- Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, 36(2), 89–101. https://doi.org/10.1207/S15326985EP3602_4
- Park, H. (2008). Home literacy environments and children's reading performance: A comparative study of 25 countries. *Educational Research and Evaluation*, 14(6), 489–505.
- Praetorius, A. K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of three basic dimensions. *ZDM*, 50(3), 407–426.
- Riddell, A. (2008). *Factors influencing educational quality and effectiveness in developing countries: A review of research*. GTZ.
- Rosenshine, B., & Stevens, R. (1986). Teaching functions. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 376–391) Macmillan.
- Savage, J. (2012). Improving quality in education: Dynamic approaches to school improvement. *Journal of Educational Administration and History*, 44(4), 396–398. <https://doi.org/10.1080/00220620.2012.718673>
- Scheerens, J. (2013). The use of theory in school effectiveness research revisited. *School Effectiveness and School Improvement*, 24(1), 1–38. <https://doi.org/10.1080/09243453.2012.691100>
- Scheerens, J. (2016). *Educational effectiveness and ineffectiveness: A critical review of the knowledge base*. Springer. <https://doi.org/10.1007/978-94-017-7459-8>
- Schoenfeld, A. H. (1998). Toward a theory of teaching in context. *Issues in Education*, 4(1), 1–94. [https://doi.org/10.1016/S1080-9724\(99\)80076-7](https://doi.org/10.1016/S1080-9724(99)80076-7)
- Scheerens, J., & Bosker, R.J. (1997). *The foundations of educational effectiveness*. Pergamon.
- Sprinthall, N., Reiman, A., & Thies-Sprinthall, L. (1996). Teacher professional development. In J. Sikula, T. Buttery, & E. Guyton (Eds.), *Handbook of research on teacher education* (2nd ed., pp. 666–703) Macmillan.
- Stenmark, J. K. (1992). *Mathematics assessment: Myths, models, good questions and practical suggestions*. NCTM.
- Tomlinson, C. A. (2005). Grading and differentiation: Paradox or good practice? *Theory Into Practice*, 44(3), 262–269. https://doi.org/10.1207/s15430421tip4403_11
- Van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research*, 49(2), 127–152.
- Walberg, H. J. (1984). Improving the productivity of America's schools. *Educational Leadership*, 41(8), 19–27.
- Walberg, H. J. (1986). Syntheses of research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 214–229) Macmillan.
- William, D., Lee, C., Harrison, C., & Black, P. J. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles Policy and Practice*, 11(1), 49–65. <https://doi.org/10.1080/0969594042000208994>

Leonidas Kyriakides is Professor of Educational Research and Evaluation at the Department of Education of the University of Cyprus, Cyprus. His main research interests are in the area of school effectiveness and school improvement and especially in modelling the dynamic nature of educational effectiveness and in using research to promote quality and equity in education. Leonidas acted as chair of the AERA SIG on School Effectiveness and Improvement and of the EARLI SIG on Educational Effectiveness. email: kyriakid@ucy.ac.cy

Anastasia Panayiotou is a post-doctoral researcher at the Department of Education of the University of Cyprus. During the last ten years, she has participated in several international and national projects aimed at identifying factors that can promote educational effectiveness and was also involved in in-service teacher professional development programs. Her main research interests are in the area of educational effectiveness and focus on teacher evaluation and professional development. email: anas.panayiotou@gmail.com

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 3

Teacher and Teaching Behaviour and Student Motivational Outcomes: Critical Reflections on the Knowledge Base and on Future Research



Marie-Christine Opdenakker

Abstract In this chapter, (a selection of) current conceptualizations, theories, measurements, and instruments of (quality of) teacher and teaching behaviour from a variety of perspectives, namely educational and teacher effectiveness research, learning environments research and research on motivational teaching are discussed. Furthermore, attention is paid to topics such as the dimensionality of teacher and teaching behaviour, and of teaching skills, as well as the existence of teaching styles and stages in teaching skill development. In addition, context, antecedents, informant as well as (in)stability issues concerning teacher and teaching behaviour are addressed. Relevant empirical findings concerning the already mentioned issues as well as empirical findings with regard to teacher and teaching effectiveness in relation to student motivational outcomes are reviewed and discussed. Attention is paid to unique and joint effects of teacher and teaching behaviour dimensions and relative sizes of effects. In addition, differential effectiveness of teacher and teaching behaviour in relation to student background characteristics such as gender, social-economic status, cognitive ability, race and ethnicity, and prior engagement is discussed. The chapter ends with conclusions, reflections, implications and suggestions for future research directions and practice related to effective teacher and teaching behaviour based on the findings discussed before.

Keywords Teacher behaviour · Motivation · Instruments · Differential effects · Stability

M.-C. Opdenakker (✉)
University of Humanistic Studies, Utrecht, The Netherlands
University of Groningen, Groningen, The Netherlands
e-mail: m.c.j.l.opdenakker@rug.nl

© The Author(s) 2023
R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_3

1 Introduction

How can students be motivated and stay motivated and what influences can teachers have on their students' motivation and learning? These questions have been triggering teachers, teacher trainers and researchers for many decades. After all, it is a well-known fact that learning takes more easily place when students are motivated (Stipek, 1988) and this is also recognized in models of learning (e.g., Illeris, 2009). Interest in the effects that teachers and, in particular, their behaviour may have on students can be found in various domains of educational research such as educational and teacher effectiveness research, learning environments research and research in the domains of educational, developmental and motivational psychology. In all these domains, conceptualizations of teacher behaviour exist as well as ideas on what constitutes a good, successful, or effective teacher. This led to the construction (and refinement) of instruments to measure relevant aspects of teacher behaviour and to the formulation of several theories. Because the domains already mentioned have different backgrounds and frameworks, and operated in the past rather independently from each other, it is interesting and important to compare their conceptualizations, measurements and instruments of teacher and teaching behaviour¹ and their findings in relation to student motivational outcomes. This operation includes looking for convergence and divergence on these topics across these domains and also addressing the dimensionality of teacher quality and effectiveness, the existence of teaching styles and stages in teaching skill development, and exploring context, informant and stability issues concerning teacher and teaching behaviour). It can enlarge our knowledge on and insights in the way in which teachers may and can have an impact on their students' motivation and how teachers' behaviour and its effect on student motivational outcomes can be optimally investigated. In this chapter, these topics will be critically addressed and substantiated with empirical findings, and findings from the mentioned domains regarding teacher and teaching effectiveness in relation to student motivational outcomes will be discussed.

¹In this chapter the terms teacher and teaching behaviour are used. In fact, teacher behaviour is a broader concept than teaching behaviour and it can include teaching behaviour. Nevertheless, it was opted to mention teaching behaviour in addition to teacher behaviour because it depends on the theoretical framework which concept is used in publications (and I wanted to stay as close as possible to the concepts used by authors in publications) and because it is informative to know if or that teaching behaviours of teachers are addressed in theoretical frameworks, conceptualizations and other relevant topics discussed in this chapter.

2 Conceptualizations of Teacher and Teaching Behaviour from a Variety of Perspectives

It is striking how many different terms are used in the literature to refer to classroom processes or practices and behaviour of teachers who appear to be good, successful, or effective in their teaching (Leon et al., 2017). For example, terms like teaching quality (Allen et al., 2011; Fauth et al., 2014; Leon et al., 2017), quality of teaching (Hattie, 2009; Teddlie et al., 2006), instructional quality (Klieme et al., 2009; Lipowsky et al., 2009; Rjosk et al., 2014), quality of instruction (Creemers, 1994; Opdenakker, 2020), teaching effectiveness (Hamre et al., 2013; Marsh & Roche, 1997; Seidel & Shavelson, 2007;), effective teaching (Campbell et al., 2004; Creemers, 1994; Muijs & Reynolds, 2011), teacher effectiveness (Campbell et al., 2004; Doyle, 1977; Kyriakides et al., 2020; Muijs et al., 2014) and classroom quality (McLean & Connor, 2015) are used. In addition, in some studies reference is made to effective teaching styles (Campbell et al., 2004; Opdenakker & Van Damme, 2006; Wentzel, 2002), instructional style (Jang et al., 2010), quality of teacher-student interactions (Hafen et al., 2015; Hamre & Pianta, 2010), and effective classroom management (Arens et al., 2015). Furthermore, some of these terms have a broader and others a narrower meaning, and sometimes it depends on who is using the term. A good example is quality of teaching (see e.g., Teddlie et al., 2006), which is often used with a narrower meaning than teacher effectiveness (Campbell et al., 2004; Muijs et al., 2014; Teddlie et al., 2006). For example, teacher effectiveness is defined by Campbell et al. (2004) as ‘the power to realize socially valued objectives agreed for teachers’ work, especially, but not exclusively, the work concerned with enabling students to learn’ (Campbell et al., 2004, p. 4). It refers to the impact of classroom factors such as teaching methods, teaching expectations, classroom organization and the use of classroom resources (p. 3). This is a broader definition than the definition of quality of teaching by Teddlie et al. (2006). They define quality of teaching by referring to indicators such as clarity of instruction, (demonstrating) instructional skills, promoting active learning and developing metacognitive skills in students, and (having an adequate) planning of single lessons. However, broader definitions of teaching are found as well. For example, Sykes and Wilson (2015) refer to two domains namely instruction and professional role responsibilities in their framework for competent teaching, a framework that was based on an interpretive synthesis of main and contemporary currents in the research on teaching and learning. The first domain (instruction) refers to preparing and planning for high-quality instruction, attending to relational aspects of instruction, establishing and maintaining the social and academic culture, interactive teaching, and engaging in instructional improvement. The second domain of teaching (professional role responsibilities) refers to collaborating with other professionals, working with families and communities, fulfilling ethical responsibilities, and meeting legal responsibilities. In addition, Campbell et al. (2004) mention that teacher effectiveness is (often) conceptualized too narrowly in the literature and that attention should be paid to differential teacher effectiveness which takes into account that teachers may

be more effective with some categories of students, some subjects and some teaching contexts than with others.

Moreover, a number of models and theories on effective teaching (e.g., the comprehensive model of educational effectiveness of Creemers, 1994; the dynamic model of educational effectiveness of Creemers & Kyriakides, 2008; Kyriakides et al., 2020), instructional quality (e.g., the three dimensions model of instructional quality of Klieme et al., 2009), and (need-)supportive teaching (e.g., the self-system process model of motivational development of Connell & Wellborn, 1991; the self-determination theory of Ryan & Deci, 2017; the teaching through interactions framework cf. Hafen et al., 2015; Hamre et al., 2013)² have been developed. Some of these theories focus mainly on how to achieve student learning outcomes, while others focus on more general/broader outcomes (e.g., well-functioning, development) or on non-cognitive outcomes such as motivation or motivated student behaviour in the classroom, or on a diversity of outcomes (cognitive as well as on non-cognitive outcomes). In addition, depending on the research domain, theorizing got more/less attention in the past. For example, in the domain of learning environments research, the focus has always been strongly on developing instruments, while theorizing got less attention. An exception is the theoretical work of Wubbels and colleagues on interpersonal behaviour of teachers. In the next paragraph, (teacher/teaching behaviour) factors often mentioned in the above-mentioned research domains and visible in famous, influential (current) theories/models stemming from these domains and included in a listing of findings of a state-of-the-art on teacher effectiveness research (Muijs et al., 2014) will be discussed. (For an overview of the selected theories/models/state-of-the-art, see Table 3.1).

Table 3.1 reveals that the theories/models and list in the state-of-the-art on teacher effectiveness refer to a different number of relevant factors/dimensions/domains, although three of them refer to three overarching factors. However, looking into more detail into these factors and their content, it is striking that there is much in common even though the different theories/models stem from a variety of research domains and their knowledge bases are mostly separately constructed. Another observation is that, depending on the research domain, some factors are more elaborated, which often results in more separate dimensions. In the following, the research domains with corresponding theories/models will be discussed paying attention to convergences and divergences.

Teacher effectiveness research and accompanying frameworks/theories refer, first, to the importance of structured teaching (including aspects of direct instruction) (Creemers, 1994; Klieme et al., 2009; Kyriakides et al., 2020; Muijs et al., 2014; Opdenakker, 2020; Opdenakker & Minnaert, 2011; Opdenakker & Van Damme, 2006; Teddlie et al., 2006; van de Grift, 2007). Structured teaching entails the delivery of explicit and clear instruction as well as structuring the lessons (clearly stating goals, making the structure of the lesson explicit, paying attention to main ideas of the lesson) and also entails elements of direct instruction such as giving an orientation on the learning content, offering explicit strategy instruction and

²Hamre et al. (2013) also use the term teacher effectiveness.

Table 3.1 Overview of factors, dimensions, and outcomes in a selection of theories/models of teacher effectiveness, instructional quality and effective teaching and mentioned in a state-of-the-art review of teacher effectiveness

Theory/model	
Author(s)	Creemers (1994) (selection – teacher behaviour)
Name of theory/model/framework	Comprehensive model of educational effectiveness
Domains/factors/dimensions	Management/orderly and quiet atmosphere High expectations Homework Clear goal setting (restricted set, emphasis on basic skills, on cognitive learning and transfer) Structuring the content Clarity of presentation Questioning Immediate exercises Evaluation Feedback Corrective instruction
	Kyriakides et al. (2020)
	Dynamic model of educational effectiveness
	Orientation modelling Structuring Questioning Application Assessment Management of time Classroom as a learning environment For each factor 5 dimensions: Frequency, focus, stage, quality & differentiation
	Klieme et al. (2009)
	Model of instructional quality ^a
	Cognitive activation & deep content Classroom management (including clarity and structure) Supportive climate ^b
	Muijs et al. (2014)
	State-of-the-art on teacher effectiveness research
	Classroom management (related to opportunity to learn and time on task) Instruction and interaction: Structuring Classroom climate (social, businesslike, and supportive for learning) Teacher expectations Attention needed for self-regulated learning
	Hafen et al. (2015)
	Teaching through interactions framework
	Emotional support – <i>Positive climate</i> – <i>Negative climate</i> ^c – <i>Teacher sensitivity</i> – <i>Regard for student Perspectives</i> ^d Classroom organization – <i>Behaviour management</i> – <i>Productivity</i> – <i>Instructional learning Formats</i> ^e Instructional support – <i>Content understanding</i> – <i>Analysis and inquiry</i> – <i>Quality of feedback</i> – <i>Instructional dialogue</i>
	Connell and Wellborn (1991)/Ryan and Deci (2017)
	Self-system process model of motivational development/self-determination theory
	Structure Autonomy support Teacher involvement

(continued)

Table 3.1 (continued)

Theory/model	
Author(s)	Creemers (1994) (selection – teacher behaviour)
Outcomes	Kyriakides et al. (2020) Achievement (Subject motivation) (Thinking style)
	Klieme et al. (2009) Student learning (i.e., <i>knowledge & understanding</i>) Motivation (<i>emotions & affects</i>)
	Muijs et al. (2014) Cognitive & non-cognitive (e.g., <i>self-concept, wellbeing, motivation & engagement</i>)
	Hafen et al. (2015) Social and emotional skills Engagement Regulation of behaviour Regulation of attention Cognition & learning Higher-order thinking/ Metacognition
	Connell and Wellborn (1991)/Ryan and Deci (2017) Satisfaction of the need to feel related, competent & autonomous Motivation Engagement Learning & achievement

^a This model is also known under the names of ‘Generic dimensions of teaching quality’ and ‘German framework of three basic dimensions’

^b This refers to supportive teacher-student relationships, positive and constructive teacher feedback, a positive approach to student errors and misconceptions, individual learner support, caring teacher behaviour (Klieme et al., 2009, p. 141). Reference is made to the fulfillment of students’ basic psychological needs as mentioned in the self-determination theory

^c This factor belongs to classroom organization in the framework for secondary education

^d Within the developmental and motivational literature, this dimension belongs to the construct of autonomy support (e.g., Anderman & Midgley, 1998; Skinner & Belmont, 1993; Hamre & Pianta, 2010)

^e This factor belongs to instructional support in the framework for secondary education

guided practice etc. There is overlap with the concept of clarity of instruction often mentioned in learning environments research³ (den Brok et al., 2006), although clarity of instruction is often more narrowly conceptualized.

In addition, teacher effectiveness research also mentions the importance of good classroom management (Klieme et al., 2009; Kyriakides et al., 2020; Muijs et al., 2014; Opdenakker, 2020; Opdenakker & Minnaert, 2011; Teddlie et al., 2006; van de Grift, 2007), and teacher behaviour that stimulates a positive relational and learning climate in the classroom (Klieme et al., 2009; Kyriakides et al., 2020; Muijs et al., 2014; Opdenakker, 2020; Teddlie et al., 2006). A positive relational climate is characterized by good and frequent teacher-student interactions and good relationships characterized by mutual respect, trust and interest in each other. A good learning climate refers to a class climate that is supportive and conducive to learning (van de Grift, 2007). In some teaching effectiveness studies the importance of the teacher as a helpful person is stressed (Opdenakker & Minnaert, 2011; Teddlie et al., 2006). The mentioned concepts also show resemblance with factors referred to as important in learning environments research, namely of classroom management (see e.g., Back et al., 2016; den Brok et al., 2006; Fraser, 2012) and teachers' interpersonal behaviour referring to proximity/communion (see e.g., den Brok et al., 2004, 2006; Wubbels, 2019; Wubbels & Brekelmans, 2005). Also, the importance of teachers' role in creating a positive psychosocial climate in the classroom and the importance of teacher involvement (Fraser, 2012) is emphasized in learning environments research.

Moreover, teacher effectiveness research points to the importance of making expectations about learning (and corresponding evaluation) explicit, and of having high and realistic student expectations as a teacher (Hattie, 2009; Muijs et al., 2014; van de Grift, 2007). The importance of providing positive and constructive feedback to students is stressed as well (Hattie, 2009; Klieme et al., 2009; Kyriakides et al., 2020; Muijs et al., 2014). Slavin (2021) points out the relevance of intentionally/ (purposeful) teaching. Furthermore, teacher behaviour in line with constructivist concepts of learning (that stimulates active student involvement in their own learning and the development of metacognitive skills) is, rather recently, receiving attention as effectiveness enhancing teacher behaviour as well (Klieme et al., 2009; Kyriakides et al., 2020; Muijs et al., 2014; Opdenakker, 2020; Teddlie et al., 2006). Lastly, teacher effectiveness research refers to the importance of offering adaptive education/instruction and differentiation opportunities (Creemers & Kyriakides, 2008; Kyriakides et al., 2020).

Theories and literature on educational, developmental and motivation psychology refer to the same kind of factors referring to providing structure, stimulation of self-regulated learning/student participation, climate, and classroom management.

³The instruments that were constructed within the learning environments research tradition to make the characteristics of the learning environments visible and to get an impression of the quality of the psychosocial climate the teachers had created in their classrooms, deliver a good illustration of this emphasis. For an overview and description of the most famous instruments, see Fraser (2012, 2019).

See for example the Teaching through interactions framework (TTI) (and research based on this framework). In this framework (see Hafen et al., 2015), which combines developmental theory with classroom practices, reference is made to three overarching factors namely emotional support (which refers to the climate in classes, teacher sensitivity and teacher's regard for student perspectives), classroom organization (which refers to, among others, behaviour management and productivity in relation to time), and instructional support (which is indicated by, among others, teachers' approaches to help students with subject matter comprehension, facilitation of higher-level thinking skill use and metacognition, quality of teachers' feedback and encouragement of students' participation, and purposeful use of dialogue-structured, cumulative questioning and discussion to facilitate students' understanding of the subject matter). The resemblance of the first factor with the already mentioned climate factor and teacher involvement in other frameworks, the second factor with classroom management, and the third factor with providing structure and the stimulation of self-regulation and participation is clear.

Related factors are visible in theories/models focusing on supporting students' motivation and engagement such as the self-determination theory (SDT; Ryan & Deci, 2000, 2002, 2017) and the self-system process model of motivational development (Connell & Wellborn, 1991), a model grounded in self-determination theory. In this model/theory it is stressed that every person requires the fulfillment of three fundamental innate psychological needs in order to function well, to flourish, to be and to stay motivated, and to experience psychological growth and well-being (Ryan & Deci, 2000). These needs are the need to feel competent, to feel autonomous and to feel related. Three (need-supportive) factors are mentioned that can satisfy these needs, namely structure, autonomy support and teacher involvement.

Structure refers to the creation of a supportive well-structured environment and includes offering optimal challenges, instrumental help and support, and positive and rich efficacy supportive feedback to students. It also includes adjusting teaching strategies to the level of the student (Ryan & Deci, 2020). In addition, it refers to the amount of information that is available in the context about how to effectively achieve desired outcomes (Connell & Wellborn, 1991; Skinner & Belmont, 1993). Structure can be provided by clearly communicating expectations and goals towards students and by responding contingently, consistently, and predictably to them. It entails the provision of clear and consistent guidelines and rules in the classroom. Structure is considered to play an important role in the fulfillment of the need to feel competent (Ryan & Deci, 2020) and is important to promote motivation and engaged behaviour (Ryan & Deci, 2002). Providing structure may not be confused with controlling teacher behaviour which pressures students to think, feel or behave in a certain way or which pressures to achieve. The 'opposite' of structure is chaos, uncertainty, and inconsistency.⁴

⁴Recently, SDT researchers have begun to see and study these need-supportive and their need-thwarting "opposites" as separate dimensions (Opdenakker, 2021; Reeve et al., 2014). Furthermore, it is recognized that little support for the needs will lead to experiences of low/deprived need satisfaction, while a more direct thwarting of individuals' needs lead to need frustration experiences (Ryan & Deci, 2017).

Autonomy support refers to supporting students to take ownership and initiative of their schoolwork (Ryan & Deci, 2020). It can be promoted and supported by providing students meaningful choices and tasks and by allowing them latitude in their learning activities, by making connections between school activities and students' interests and by offering students a rationale for tasks and learning activities that must be done. It also entails attempts to understand, acknowledge, respect, and where possible, be responsive to the perspective of students, to give them a voice and to use informational language (Ryan & Deci, 2017). For fostering autonomy, the absence of controls and pressures and, also, of external rewards is important. Autonomy support is seen as promoting not only the satisfaction of the need to feel autonomous but contributes also to the satisfaction of the need to feel related and when it occurs along with structure, the satisfaction of competence is promoted as well. In addition, in respecting autonomy and advocating for its support, which entails, as mentioned before, respecting and attempting to appreciate the perspective of each student as well as his/her unique challenges, the importance of differences between students is acknowledged as well (Ryan & Deci, 2020). The 'opposite' of being autonomy supported is being coerced and feeling controlled (Connell & Wellborn, 1991; Skinner & Belmont, 1993). Controlling teachers are more oriented to pressure students with regard to their thinking, feeling or behaving and are not responsive to student perspectives.

The third factor, teacher involvement, is of particular importance to fulfill students' need of relatedness and refers to creating a caring, supporting and respectful environment (Ryan & Deci, 2020). It entails expressing warmth and affection towards students, enjoying interactions with them, taking time for them, and being attuned and dedicate resources to them. Involvement refers to the quality of the interpersonal relationship with teachers and peers. The 'opposite' of involvement is rejection or neglect.

The structure factor resembles structure and classroom management factors in other frameworks, while the teacher involvement factor is familiar with (relational) climate and emotional support⁵ factors in other frameworks. The autonomy support factor has connections with factors referring to the stimulation of students' self-regulation and to teacher actions in line with constructive ideas of learning mentioned in other frameworks.

In general, it can be concluded that all these frameworks and theories mentioned and discussed in the preceding pages include combinations of factors/dimensions that were associated with different research domains in the earlier days. For example, a strong focus on instruction and instructional context is characteristic for educational research, while social dynamics of and within the class has always got much attention in developmental and learning environments research (Hamre & Pianta, 2010). Classroom management and organization has always been a factor that was highly focused on in research on teaching and teacher training, learning

⁵This familiarity between teacher involvement of the SDT and emotional support of the TTI is also recognized in Virtanen et al. (2018).

environments research (Hamre & Pianta, 2010), and educational psychology (Emmer & Strough, 2001). Overlooking the dimensions of the discussed frameworks and theories, they all have a rather broad and holistic approach to and vision on (the quality of) teacher behaviour. However, it is also clear that there are some differences regarding the degree to which the dimensions are elaborated. For example, it is obvious that instruction is quite elaborated within the models and frameworks related to teacher effectiveness research, while teachers' role in creating a positive psychosocial classroom climate and offering emotional support is less well elaborated, in particular, in the oldest ones. In other frameworks e.g., the TTI or Need-supportive teaching framework, these dimensions are more equally elaborated.

3 Measurements and Instruments of Teacher and Teaching Behaviour

In each of the mentioned domains of research, instruments for the (reliable and valid) measurement of teacher/teaching behaviour were developed in line with theoretical perspectives, models, and knowledge bases. A comparison of these instruments reveals that they differ regarding the type of informants (teachers – self-report, student perspectives, observers, consultants/administrators), the kind of data collection method used (questionnaires, observation instruments, vignettes, etc.), and the intended educational level (preprimary, primary, secondary education). In the early developing phases of the instruments, the choices made in this respect were the logical consequence of the research traditions in the domains concerned and were often conceived as generic instruments. Later, additions were made to some of the existing instruments. For example, observation variants were added to questionnaires tapping student perceptions (or vice versa), different forms were made to map not only the current perception of teacher's classroom behaviour/classroom environment, but also the ideal (i.e., preferred teacher behaviour/classroom environment) or the expected teacher behaviour/classroom environment. Sometimes, adaptations for other educational levels than the original were made as well. One of the most known and wide-spread used instruments are the *CLASS* [Classroom Assessment Scoring System] instrument (Pianta & Hamre, 2009; Pianta et al., 2012) stemming from the domain of developmental and educational psychology), the *WIHIC* [What Is Happening In this Class] from the domain of learning environments research⁶ (Fraser et al., 1996), the *ICALT* [International Comparative Analysis of Learning and Teaching] (van de Grift, 2007), the *ISTOF* [International System for Teacher Observation and Feedback] instrument (Muijs et al., 2018; Opdenakker & Minnaert, 2011; Teddlie et al., 2006), both stemming from educational and teacher

⁶Another famous instrument is the CES (Moos & Trickett, 1974). Due to word constraints and because the CES is older than the WIHIC, this instrument was not included in this review.

effectiveness research, and the *TASC* [Teacher As a Social Context] (Belmont et al., 1992), which is based on elaborations of the self-determination theory/self-system processes model of motivational development (Ryan & Deci, 2017, 2020; Connell & Wellborn, 1991).

A comparison of these instruments reveals that, in line with the findings about the theoretical/knowledge base foundations of these instruments, the instruments share overlapping concepts and characteristics that are recognized as effective teaching behaviour in teacher effectiveness research (see Table 3.2). For a description and discussion of these instruments, see the Appendix.

4 Dimensionality, Stability and Best Informants of Teacher and Teaching Behaviour

4.1 Dimensionality of Teacher and Teaching Behaviour

An important question is how the mentioned dimensions/factors/domains of the instruments described in the preceding section and the appendix should be considered. Do they refer to a one-dimensional, multidimensional or multifaced conceptualization of teaching and teacher behaviour? What evidence does validation research deliver about the theoretical conceptualizations?

In general, all the dimensions/factors/domains distinguished in the instruments are, from a theoretical point of view, considered as unique contributors to teaching and a lot of validation studies found evidence for the multidimensionality of teacher behaviour.⁷ For example, a variety of studies (e.g., Allen et al., 2013; Hafen et al., 2015; Hamre et al., 2013; Virtanen et al., 2018) found evidence for the three-domain latent structure of the CLASS/CLASS-S instrument. In each of the studies, a three-factor solution (in confirmatory factor analysis) had a better fit compared to one- or two-factor solutions. The studies referred to a variety of classroom settings (ranging from preschool to high school) and to teaching in a variety of countries. Comparable findings providing evidence for the multidimensionality of teacher behaviour/teaching were found with regard to the WIHIC (e.g., Aldridge & Fraser, 2000; Dorman, 2003), the TASC (e.g., Opdenakker, 2014; Sierens et al., 2009⁸; Vansteenkiste et al., 2012⁹) and dimensions related to need-supportive teaching (Jang et al., 2010¹⁰), the

⁷However, there are also a few exceptions related to the CLASS as well as the ISTOF instrument. For a discussion of the first, see Virtanen et al. (2018), and for the second, see Muijs et al. (2018).

⁸In this study, only autonomy support and structure were included. Confirmatory factor analysis indicated a significantly better fit for the two-factor model compared to the one-factor model.

⁹In this study, a short version with an adaptation of the dimension 'structure' was used.

¹⁰Jang et al. (2010) distinguished, in an observation instrument, between autonomy support and structure and found evidence based on confirmatory factor analysis that a two-factor model had a significant better fit than a one-factor model. However, they also explored how both dimensions relate to each other (antagonistic, curvilinear, independent) and found that both relate in a linear way.

Table 3.2 Overview of a selection of instruments tapping teacher behaviour with corresponding factors/dimensions of teacher behaviour

Name	CLASS (classroom assessment scoring system)	WIHIC (what is happening in this class?) ^a	Instrument			ISTOF - questionnaire ^b	TASC (teacher as a social context)
			ICALT (international comparative analysis of learning and teaching)	ISTOF (international system of teacher observation and feedback) -observation	ISTOF (international system of teacher observation and feedback) -observation		
Factors/ dimensions/ domains/ components (and subdimensions)	Emotional support – <i>Positive climate</i> – <i>Negative climate</i> ^c – <i>Teacher sensitivity</i> – <i>Regard for adolescent perspectives</i> ^d	Student cohesiveness Cooperation Equity Teacher support ^e Task orientation ^f	Learning climate (safe and stimulating)	Classroom climate	(see helpful)	Teacher involvement	
	Classroom organization – <i>Behaviour management</i> – <i>Productivity</i> – <i>Instructional learning formats</i> ^e		Classroom management (efficient)	Classroom management	Manager/organizer of classroom activities	Structure	
	Instructional support – <i>Content understanding</i> – <i>Analysis and inquiry</i> – <i>Quality of feedback</i> – <i>Instructional dialogue</i>	Teacher support ^g Investigation	Clarity of instruction (clear instruction) Teaching learning strategies	Clarity of instruction Instructional skills	Helpful and good instructor	Autonomy support	
		Involvement	Activating teaching	Promoting active learning and developing metacognitive skills (see also instructional skills)	Teacher as promotor of active learning and differentiation		

			Differentiation/ adaptive teaching (adaptation of teaching)	Differentiation and inclusion Assessment and evaluation	(see differentiation)
--	--	--	---	--	-----------------------

^a The original version of this instrument also included the dimension autonomy/independence (see Fraser et al., 1996)
^b ISTOF questionnaire (see Opdenakker & Minnaert, 2011)

^c This factor belongs to classroom organization in the instrument for secondary education (CLASS-S; Pianta et al., 2012)

^d Within the developmental and motivational literature, this dimension belongs to the construct of autonomy support (e.g., Anderman & Midgley, 1998; Skinner & Belmont, 1993; Hamre & Pianta, 2010)

^e This factor belongs to instructional support in the instrument for secondary education (CLASS-S; Pianta et al., 2012)

^f This dimension refers to the task orientation of the student and includes, for example, if it is important for them to get a certain amount of work done or to understand the work. It also refers to knowing the goals for the class

^g This dimension includes aspects of emotional as well as instructional support

ISTOF (student questionnaire: Opdenakker & Minnaert, 2011; observation instrument: for a review, see Muijs et al., 2018) and the ICALT (e.g., Maulana et al., 2017, 2021; Maulana & Helms-Lorenz, 2016; van de Grift et al., 2011¹¹).

In addition, regarding some conceptualizations/instruments, evidence was found for the usefulness of a conceptualization in terms of a circumplex model which offered the opportunity to combine dimensions in order to distinguish between teaching styles. A well-known use of the circumplex model is related to dimensions of the Questionnaire on Teacher Interaction, an instrument rooted in learning environments research (Brekelmans et al., 2011). Recently such an approach was successfully adopted as well by Aelterman et al. (2019) using two (of the three)¹² dimensions of need-supportive teaching in line with the SDT framework namely autonomy support and structure. Aelterman et al. (2019) collected self-reports from Belgian secondary school teachers and students using the vignette-based Situations-in-School Questionnaire and applied multidimensional scaling analyses. This resulted in a two-dimensional configuration forming a circumplex with eight subareas, namely participative and attuning, guiding and clarifying, demanding and domineering, and abandoning and awaiting. The correlations between these subareas and various outcome variables followed the expected sinusoid pattern.

Furthermore, although the instruments discussed before can differentiate between the different factors/dimensions/domains and validation studies deliver evidence for the existence of these different factors/dimensions/domains, there are also indications in the literature of positive associations between the factors/dimensions/domains. This could lead to some confusion regarding how the relationship between the dimensions should be conceptualized. Den Brok et al. (2019), reviewing instruments rooted in learning environments research, mention that correlations between dimensions of these instruments often range between 0.20 and 0.60. This indicates some overlap as well as idiosyncrasy. Regarding other instruments rooted in different theoretical frameworks, similar findings are reported. For example, Jang et al. (2010) mention, based on observation measures within the SDT framework, a positive correlation between autonomy support and structure ($r = 0.60$). Also, Sierens et al. (2009) found that autonomy support and structure (of math/Dutch language/educational science teachers as perceived by their students from grade 11–12 academic track classes) is correlated ($r = 0.67$), which is confirmed by Lietaert et al. (2015) doing research in grade-7 Dutch language general and vocational track classes ($r = 0.71$), and by Hospel and Galand (2016) in French language grade-9 vocational and general classes in the French-speaking part of Belgium ($r = 0.60$).

¹¹In this study, primary teachers of the Netherlands, Flanders (Belgium), Germany, Slovakia, Croatia, and Scotland were observed.

¹²The third dimension, namely teacher involvement, which relates to relatedness support, should be studied as well in relation to the circumplex model, since need-supportive teaching relates to three dimensions in order to fulfill the three basic psychological needs of feeling autonomous, competent and related. This view is underscored by Vansteenkiste et al. (2020).

Confirmation is also found in the study of Vansteenkiste et al. (2012)¹³ who report a significant correlation ($r = 0.54$) between autonomy support and clear expectations, a subdimension of structure based on research in grade 7–12 mainly general track classes. In addition, Vansteenkiste et al. (2012) found based on cluster analysis evidence for four teaching configurations¹⁴ of which two referred to scoring high or low on both dimensions and two configurations scoring high on one of the two dimensions. Furthermore, Lietaert et al. (2015) reported somewhat lower, but significant, correlations between teacher involvement and autonomy support and structure (respectively $r = 0.58$ and $r = 0.59$).

In addition, regarding the dimensions of the CLASS/CLASS-S instrument similar findings are reported (cf. Pianta et al., 2012). For example, Pöysä et al. (2019) mention correlations between 0.52 and 0.62 in their study on grade-7 Finnish mathematics and language art classes ($r = 0.52$ between instructional support and classroom organization, $r = 0.62$ between instructional support and emotional support, and $r = 0.61$ between emotional support and classroom organization), while Virtanen et al. (2015) report correlations between 0.37 and 0.75 based on observations in Finnish grade-7 literacy, history and civics, science and home economics classes ($r = 0.37$ between instructional support and classroom organization, $r = 0.75$ between instructional support and emotional support, and $r = 0.48$ between emotional support and classroom organization). Reyes et al. (2012) mention comparable correlations related to fifth/sixth-grade classes: $r = 0.57$ between instructional support and classroom organization, $r = 0.68$ between instructional support and emotional support, and $r = 0.60$ between emotional support and classroom organization.

Also, regarding the dimensions of the ICALT observation instrument, clear evidence for associations between dimensions is found. Van de Grift et al. (2011) report correlations¹⁵ between 0.55 and 0.92 with an average correlation of 0.75. Adaptive teaching has the lowest correlations with other dimensions (average correlation: 0.64) and the climate dimension the second lowest (average correlation: 0.70). The reported correlations are quite high in comparison with the mentioned ones of other instruments. One of the reasons could be that several dimensions of the ICALT refer to teacher behaviour related to instruction. Regarding the ICALT, also the one-dimensionality of the scale was explored and evidence for it was found in several studies (e.g., van de Grift et al., 2011; van de Grift et al., 2014; Maulana

¹³They used the autonomy support dimension of the short version of the TASC (Dutch translation). For the dimension 'clear expectations', the 'clarity of expectations' of the Structure scale of the TASC (Belmont et al., 1988) was used as a source of inspiration. This scale was elaborated by (formulating additional) items on expectations regarding (1) the learning material and tests, and (2) desirable behaviour in class.

¹⁴To some degree the configurations deliver evidence for the distinctness of the dimensions, although also evidence is found for a positive relation between them (since two out of four configurations refer to scoring in the same way on both dimensions). Moreover, the authors mention that they did not find strong evidence for unique correlates of both dimensions, albeit some relevant exceptions were found as well. Yet, several exceptions deserve being discussed.

¹⁵The reported correlations are LISREL based ϕ -coefficients.

et al., 2021). Furthermore, evidence was found for a systematic hierarchy in the difficulty level of teaching activities ranging from more basic (the creation of a safe and stimulating climate, efficient classroom organization and management, the provision of clear and structured instruction) to more complex (activating teaching, adaptive teaching, and teaching learning strategies) (van de Grift et al., 2011, 2014; van der Lans et al., 2018). This hierarchy is in line with Fuller's theory on the development of teachers' stages of concern (Fuller, 1969) and seems to be in line with ideas that novice teachers may need to reach a minimum level of competency in classroom management skills before they are able to develop in other areas of instruction (Emmer & Strough, 2001).

Regarding the ISTOF student questionnaire, an average correlation of 0.44 was found between factors indicating a weak-to-moderate association ($r = 0.25$ between 'teacher as promoter of active learning and differentiation' and 'classroom management', $r = 0.40$ between 'teacher as a helpful and good instructor', $r = 0.68$ between 'teacher as a helpful and good instructor' and 'teacher as promoter of active learning and differentiation' (Opdenakker & Minnaert, 2011). There are also indications of positive associations between the dimensions of the ISTOF observation instrument (for a discussion, see Muijs et al., 2018).

In general, it seems to be that the (overarching) dimensions measured with the instruments must be seen as complementary and (often) uniquely predictive of student outcomes, rather than as separate and independent of each other (Jang et al., 2010), and that the dimensions referring to instruction (and classroom organization and management) seem to refer to an overarching dimension referring to teacher activities with a different level of difficulty. This line of thought agrees with findings of Malmberg et al. (2010) who followed teachers from their last year of teacher education into their first 2 years of teaching practice and found different patterns of evolutions with regard the three dimensions of the CLASS-S (classroom and management skills, instructional support and emotional support). These findings call for considering multiple dimensions/domains rather than an overall indication when examining teaching, teaching quality, teacher effectiveness and teacher development.

4.2 *Stability of Teacher and Teaching Behaviour*

An important question, also from the perspective of obtaining good measurements of the quality of teaching and teacher behaviour, is if teaching and teacher behaviour is stable across lessons and time.

In general, not many studies have addressed this topic and in the few studies addressing (in)stability of teacher behaviour during a school year evidence is found for (small to large) changes and for, on average, mostly declining trends in the quality of teaching and student learning environment experiences from start to the end of the school year. For example, Maulana et al. (2016) reported declines in (student perceptions of) instructional behaviours (clarity and classroom management) and Opdenakker and Maulana (2010) found declines in structure, autonomy support,

and, to a lesser extent, also decreases in teacher involvement in secondary education in the Netherlands. Also, Maulana et al. (2013) found evidence for a decrease in observed teacher involvement in secondary education. In line with these studies, (small) declines in the quality of interpersonal behaviour were found in secondary education (e.g., Mainhard et al., 2011; Opendakker et al., 2012; the Netherlands) and regarding teacher involvement in primary education (Skinner & Belmont, 1993; New York). In contrast, research in secondary education in Indonesia revealed evidence for increasing quality during the school year (student perceptions) regarding involvement, structure, and autonomy support (Maulana & Opendakker, 2014) and regarding interpersonal teacher behaviour (proximity and influence) (Maulana et al., 2014). A mixed picture is visible in the study of Stroet et al. (2015). They found clear decreases of observed autonomy support and teacher involvement, and a small increase in structure in prevocational classes in the Netherlands. In all studies using multilevel growth curve modelling, evidence for differences between classes/teachers regarding the trajectories were reported as well indicating deviations from the average trend.

4.3 Best Informants of Teacher and Teaching Behaviour

Scholars in learning environment and motivation research often stress the importance of tapping students' perceptions of teachers' teaching behaviour (e.g., den Brok et al., 2005; Fraser et al., 2021; Hamre & Pianta, 2010; Ryan & Deci, 2020) and several studies revealed evidence that students' experiences of their teachers' teaching are valuable and can be reliably measured (Fauth et al., 2014; Kunter & Baumert, 2006). In addition, Kulik (2001) concludes in his review study on the validity of student ratings that student ratings have high validity (strong correlation with classroom observations and expert observations) and Cipriano et al. (2019) found evidence of agreements between primary school students of the same class regarding perceptions of teacher support: perceived teacher support at class level was significantly associated with individual student perceptions of teacher support.

Teacher questionnaires are also used, especially in large scale studies, to receive information on teachers' behaviour and the characteristics of the learning environments they create in their classes (Kunter & Baumert, 2006). Some studies addressed the agreement between student and teacher ratings. In general, these studies report weak to moderate correlations (see for example, Cipriano et al. (2019) regarding perceptions of teacher support). Studies comparing student and observer ratings refer, broadly spoken, to moderate associations (Kunter & Baumert, 2006).

Furthermore, student perceptions of their teachers' behaviour and learning environment experiences are often stronger associated with student outcomes (e.g., academic achievement or motivational outcomes) than teachers' self-report about their own teaching (Van Damme et al., 2004) or ratings of external observers (De Jong & Westerhof, 2001; Maulana & Helms-Lorenz, 2016).

Hamre and Pianta (2010) addressed the importance and advantages of observational measures focused on teaching quality and stressed that these measures are better than measuring discrete teaching behaviours since these measures may be more meaningful assessments of higher order organizations of teaching behaviour and ‘tend to parse the behavioral stream into more contextually and situationally sensitive “chunks” (p. 34).

Kunter and Baumert (2006) mention that all informants (students, teacher, observers) can have their own biases and that discrepancy between the mentioned informants can also be viewed from another perspective, namely that they can reflect perspective-specific validities. Based on their study, in which they compared student and teacher ratings of instruction, they concluded that student and teacher ratings were best suited to tapping different aspects of the learning environment. This is in line with Clausen (2002) who found, examining whether the perspectives of the three types of informants could be subsumed in a common model of instructional quality, that the data were best replicated by introducing three method factors, indicating that students, observers, and teachers tend to perceive instruction in specific ways. In addition, the method factor for students’ perceptions of instruction, showed that, although students were able to distinguish between diverse instructional aspects, their evaluation of the teacher was also shaped by a generally positive or negative attitude towards their teacher. Furthermore, Brekelmans et al. (2011) found, when examining if students and teachers use a similar frame of reference when thinking about how a teacher relates to students, that although they use a similar framework, they do not agree on the amount of teacher control/influence and affiliation/proximity in a particular class. We agree with Kunter and Baumert (2006, p. 244) that *‘because various methods have particular strengths for assessing different instructional features in research on classroom processes ... great care [should] be taken in choosing a data source appropriate for the construct to be measured.’*

5 Teacher and Teaching Effectiveness in Relation to Student Motivational Outcomes

In general, it can be stated that there is much evidence for the importance of the previously mentioned dimensions in relation to students’ learning and development. This is not surprising since authors of the instruments often explicitly mention that their instrument and underlying framework, model or theory is based on or contains, at least partly, dimensions and/or scales that have been shown in previous studies to be significant predictors of student outcomes (see e.g., Fraser et al., 1996; Hamre & Pianta, 2010; van de Grift, 2007).

However, since motivation and engagement are often seen as antecedents for learning, achievement and development, it is of great importance to explore whether the dimensions in line with the discussed frameworks and instruments are associated with motivational outcomes. Motivational outcomes refer in this review to motivation (autonomous, controlled, extrinsic, intrinsic), engagement, effort, and

motivational attitudes (e.g., interest, enjoyment, pleasure, task value, subject attitude).

To find relevant empirical studies, Web of Science, PsycINFO and Google Scholar were searched (1990–2021). Studies had to address a motivational outcome (see previous paragraph, or mention ‘motivation’/‘motivational outcome’) and refer to teaching, teacher/teaching/instructional quality/effectiveness/behaviour, quality of teaching, teacher support, class/classroom experiences, learning environment, teacher-student relationship(s) or need(–)supportive teaching/style. In addition, a reference to one of the mentioned frameworks, instruments or dimensions of the frameworks/instruments had to be included and an appropriate method of analysis (e.g., account for nested data structure if necessary) had to be used. Furthermore, recent review studies on teacher/teaching effectiveness, need-supportive teaching and quality of teacher-student relationships were consulted.

First of all, evidence was found for effects of overarching or umbrella measurements of teaching quality in line with the earlier discussed frameworks and instruments on motivational outcomes. For example, research of Klem and Connell (2004) conducted in primary and secondary education found that teacher support experiences (combining teacher involvement, structure and autonomy support items) mattered with regard to students’ engagement. Tas (2016), investigating effects of teacher support on engagement (agentic, behavioral, emotional, cognitive) in Turkish middle school science classes (grade 6 and 7) and using some of the WIHIC dimensions, among others teacher support (a combination of emotional and instructional support), found positive effects of teacher support on all engagement dimensions. In addition, the study revealed that the effect of teacher support was mediated by students’ self-efficacy (except for agentic engagement).

Also, Vandekerckhove et al. (2019), investigating the relation between weekly need-based experiences and variations (based on, among others, experiences with the teacher) and weekly academic (mal)adjustment, found positive associations between weekly variations in need satisfaction and weekly variations in engagement and autonomous motivation, and between variations in need frustration and variations in controlled motivation. In addition, research of van de Grift et al. (2011, 2014), using the teaching skill scale (RASCH scale) based on the ICALT, delivered evidence of a positive association between teachers’ teaching skill and student engagement (at class level). Van de Grift et al. (2011) reported a correlation of 0.62. Maulana and Helms-Lorenz (2016), using a student perceptions and observation version of the ICALT, also found a relationship between the teaching skill scale (observations and student perceptions) and student engagement. However, student perceptions were more strongly associated with student engagement and when both were included in a model to predict student engagement, observations were not significant anymore.

Furthermore, also regarding distinct dimensions, effects on motivational outcomes were found (see for dimensions related to SDT the review study of Stroet et al., 2013; Opdenakker, 2021). Results regarding related dimensions will be discussed together in the next pages.

5.1 Effects of Teachers' Emotional Support, Involvement, and Positive Teacher-Student Relationships

In general, clear evidence is found for positive associations between the quality of teacher-student relationships and (academic) engagement (for reviews see; Opdenakker, 2021; Roorda et al., 2011; Stroet et al., 2013). For example, Roorda et al. (2011), reviewing the influence of affective teacher-student relationships on students' academic engagement (from preschool to high school) and using a meta-analytic approach, found evidence for medium to large associations between the quality of these relationships and (academic) engagement. Also Furrer and Skinner (2003) and King (2015), investigating the relationship between students' relatedness to their teacher (and peers and parents) and students' engagement found evidence for a unique effect of relatedness to their teacher and engagement, while the studies of den Brok et al. (2004, 2005, 2010) and Opdenakker et al. (2012) revealed positive effects of teachers' proximity (a dimension of interpersonal behaviour) on students' motivational and attitudinal outcomes such as (autonomous motivation, pleasure, relevance, confidence, effort, subject attitude). Furthermore, Archambault et al. (2017) found unique effects of close teacher-student relationships on behavioral engagement in Canadian third and fourth grade primary education classes (regular and special education); however, they did not find an effect on emotional engagement. Also, the study of Lam et al. (2012), investigating the relationship between teacher (mainly emotional) support (referring to teachers at school) and student engagement (composite of emotional, behavioral, and cognitive engagement) in the lower grades of secondary education in 12 countries, revealed a significant positive association between teachers' emotional support and engagement. Likewise, Fatou and Kubiszewski (2018), studying the effect of the quality of the relationship between teachers and students (student perceptions) in grade 10–12 classes in France, found positive effects on engagement (composite of behavioral, emotional, and cognitive engagement).

Furthermore, Reyes et al. (2012), using the CLASS observation instrument, revealed that there was a positive relationship between teachers' emotional support to their class and students' engagement in fifth and sixth grade English language art classes even when controlled for the quality of class organization and teacher's instructional support¹⁶ and teacher characteristics (gender, educational attainment, teaching experience, burnout and teaching efficacy). The effects were robust for grade and gender. Furthermore, their study revealed that student engagement partially mediated the relationship between emotional support and academic

¹⁶The effects of the quality of class organization and instructional support were not significant when included in the model together with emotional support and the mentioned teacher (and non-mentioned student) characteristics. This was the case for engagement and achievement and is contrary to studies showing that, at least, instructional support matters to academic achievement (Hamre & Pianta, 2005; Mashburn et al., 2008). One possible explanation that the authors mention is that instructional support and class organization may not have fully captured because they used a CLASS version developed primarily for lower elementary classrooms.

achievement. Likewise, the Finnish study of Pöysä et al. (2019), using the CLASS-S, indicated that teacher's emotional support in grade-7 mathematics and language art classes was positively associated with students' situation-specific emotional engagement. However, they did not find significant relations with situation-specific behavioral/cognitive engagement. Virtanen et al. (2015) did not find a direct effect of emotional support on student engagement in Finnish grade 7–9 classes, however, emotional support contributed to student engagement indirectly via its effect on teachers' organizational and instructional support. Malmberg et al. (2010), also using the CLASS-S, found that observed student engagement in English classes was higher in lessons with high emotional support, classroom organization, and instructional support.

Also, other studies investigating the effects of being in emotionally supportive classrooms report positive effects on motivational outcomes such as enjoyment, interest, and engagement (e.g., Wentzel et al., 2010; You & Sharkey, 2009; Fauth et al., 2014). In addition, studies using the WIHIC in primary or secondary classes in a variety of countries found evidence for positive effects of supportive teachers on attitudinal outcomes such as enjoyment related to science, math, or language subjects (e.g., Chionh & Fraser, 2009; Telli et al., 2006; Wolf & Fraser, 2008). Other studies adopting the SDT framework and investigating associations between student perceptions of teacher involvement and motivation or academic engagement, found evidence for the importance of teacher involvement as well. For example, research of Bieg et al. (2011) shows that students' perception of teacher care in eighth grade was linked to higher intrinsic motivation in physics. Skinner and Belmont (1993) found evidence for the importance of student perceptions of teacher's involvement to emotional engagement in primary education, while Lietaert et al. (2015) and Opdenakker (2021) found positive effects on, respectively, behaviour engagement and a composite measure of behavioral and emotional engagement in secondary education (respectively in Dutch language, and EFL/math classes). Also, other work of Opdenakker, Maulana, Stroet and colleagues in the Netherlands (Maulana et al., 2013; Opdenakker, 2013, 2014; Opdenakker & Maulana, 2010; Stroet et al., 2015) indicates the importance of teacher involvement – which is important to meet students' need to feel related to significant others – in relation to student motivational outcomes and academic engagement in primary as well as in general and prevocational secondary education.

In addition, Opdenakker and Minnaert (2014) found evidence for the importance of feeling related with the teacher on primary school students' engagement. Also, the review study of Stroet et al. (2013) confirms these findings with regard to engagement and motivation, as well as their longitudinal study on associations between observed teacher involvement and motivational outcomes in grade-7 prevocational math classes (Stroet et al., 2015).

In line with this, numerous studies have found evidence for the importance of a good relational climate in classes (referring to, among others, good teacher-student relations) (For reviews, see Opdenakker, 2020; Roorda et al., 2011; Stroet et al., 2013). A few studies (e.g., Opdenakker, 2021) also paid attention to need-thwarting teacher behaviour such as teacher neglect and rejection and found negative effects

on students' engagement. Likewise, Archambault et al. (2017) found negative effects of conflictual teacher-student relationships on students' emotional engagement (for boys only). However, they did not find an effect on behavioral engagement.

Some studies also paid attention to the possibility of differential effectiveness of teachers' emotional support, involvement, and positive teacher-student relationships in relation to student (background) characteristics such as gender, socio-economical status or ethnicity. According to the academic risk hypothesis (Hamre & Pianta, 2001), teacher support in terms of an emotionally warm and caring, low-conflict teacher-student relationship is considered to be more important for students at risk (for school failure). In line with this hypothesis, the meta-analysis of Roorda et al. (2011), investigating the effect of teachers' emotional support/involvement on students' engagement, revealed that this kind of teacher behaviour was more important for boys' than for girls' engagement, indicating a higher sensitiveness of boys. Also, Furrer and Skinner (2003) and Opdenakker (2021) found support for a higher sensitiveness of boys regarding respectively perceived relatedness with the teacher, and teachers' emotional involvement and neglect/rejection.

Archambault et al. (2017) found that only boys seemed to be sensitive to conflictual teacher-student relationships regarding their emotional engagement and Fatou and Kubiszewski (2018) also found that only boys were sensitive to the quality of teacher-student relationships with regard to emotional engagement. However, when focusing on a composite of engagement, cognitive or behavioral engagement they did not find evidence for the differential effectiveness of teacher-student relationships in relation to gender. Also, other studies (e.g., Lam et al., 2012; Lietaert et al., 2015; Wang & Eccles, 2012) found no evidence for differential effectiveness regarding gender and some found that girls seemed to be more sensitive to warm and close relationships with teachers (e.g., Archambault et al., 2017). Likewise, research of Pöysä et al. (2019) suggested that girls benefited more from high emotional support than boys for their situation-specific emotional engagement.

Studies addressing differential effectiveness of teachers' emotional support related to racial or ethnic differences are rather scarce and results seem to be mixed, but when differences are found they seem to be in line with the academic risk hypothesis (Wang & Eccles, 2012; Konold et al., 2017). Den Brok et al. (2010) found no evidence for differential effects of teacher proximity on students' subject attitudes (including enjoyment, interest, and effort) related to students' ethnicity, however they found differential effects of teachers' interpersonal behaviour related to influence indicating that only students with a non-Dutch background (of the second generation) were sensitive to influence in relation to their engagement. Studies addressing differential effectiveness of the quality of teacher-student relationships in relation to the social background of students are scarce as well. Fatou and Kubiszewski (2018) studied the differential effectiveness of perceived quality of teacher-student relationships and found only evidence regarding cognitive engagement indicating that especially students with a more privileged social background were more sensitive.

5.2 *Effects of Teachers' Classroom Management and Organization*

Many studies have reported positive effects of classroom management on student academic outcomes (Seidel & Shavelson, 2007). Good classroom management helps to create good preconditions for time on task that is, in turn, crucial for students' learning and achievement (Seidel & Shavelson, 2007). An important question is whether good classroom management has also positive effects on motivational outcomes (such as engagement, intrinsic motivation for learning/working in class, and interest). Some researchers point to the possible detrimental effect it can have on students' motivational development (McCaslin & Good, 1992), since well-managed classrooms can be quite teacher-directed and are characterized by external regulation of student behaviour.

There is surprisingly little research on the effects of classroom management on motivational outcomes (Kunter et al., 2007; Korpershoek et al., 2016). Research of e.g., Klieme et al. (2009) reports positive effects of observed classroom management (based on an observation of three lessons) on students' intrinsic motivation (working interest; measured with an immediate posttest and controlled for interest in the subject mathematics at the beginning of the school year) in secondary education of schools in Germany and Switzerland. Also, Kunter et al. (2007), re-analyzing data regarding mathematics education from the German sample of the Third International Mathematics and Science Study (TIMSS, Beaton et al., 1996), found evidence for significant, but weak effects of math teachers' classroom management: (individual) students' perceptions of rule clarity and teacher monitoring were positively related to their math-related interest development. However, no (additional) effects were found for classroom management at class level. In addition, their study demonstrated that the effects of rule clarity and monitoring were partially mediated by students' experiences of autonomy and competence.

From the TTI (Teaching through interactions) framework there is some evidence for the importance of classroom organization. For example, Virtanen et al. (2015), using the CLASS-S, demonstrated a positive relation between both classroom organizational (and instructional) support and student-rated, teacher-rated, and observed general behavioral engagement among lower secondary school students in Finland. Furthermore, Pöysä et al. (2019), using the CLASS-S, found that classroom organization was positively associated with students' situation-specific behavioral/cognitive engagement in Finnish grade-7 mathematics and language art classes. However, they did not find significant relations with situation-specific emotional engagement. Also, Malmberg et al. (2010), using the CLASS-S, found evidence for the importance of the mentioned characteristic: observed student engagement was higher in lessons with high classroom organization, (and high emotional and instructional support).

Van de Grift (2007) found, using the ICALT instrument, a positive association between classroom management and observed student involvement in primary education across four European countries ($r = 0.54$). Also, van de Grift et al. (2017),

using the same instrument in a study on South Korean and Dutch secondary education teachers, reported positive associations between classroom management and observed student engagement at class level (γ -coefficients between latent dimensions and engagement at class level were respectively 0.80 and 0.79).

Also, Opdenakker and Minnaert (2011), using the student perceptions questionnaire of ISTOF, reported effects of classroom management on academic engagement in primary education in the Netherlands. However, the effect disappeared when controlled for student background characteristics (gender, nationality, language spoken at home) and prior engagement. Furthermore, Maulana et al. (2016) found small, but significant, effects of perceived classroom management in secondary education on motivational aspects such as intrinsic value and self-efficacy. However, they did not find an effect on test anxiety.

In addition, Tas et al. (2018) report that it is possible to train student teachers to improve their teaching skills and, in particular, their classroom management. They found a large effect size representing student teachers' improvement in classroom management. Furthermore, research has also established that teachers trained in classroom management principles and concepts were more likely to have engaged students compared to teachers in control groups (Emmer & Strough, 2001). In contrast, in a meta-analysis on classroom management interventions Korpershoek et al. (2016) did not find a significant effect of these interventions on student motivational outcomes. However, their results must be interpreted with caution since they were only related to six studies.

Studies addressing differential effectiveness of teachers' classroom management and organization are very scarce. Pöysä et al. (2019) investigated this in relation to student gender in secondary education and did not find evidence for differential effects on student engagement. Also, Opdenakker and Minnaert (2011), studying this in primary education, did not find evidence for differential effects related to student gender, nor did they find such effects in relation to students' prior engagement and ethnic-cultural background.

5.3 Effects of Teachers' Instruction and Instructional Support

Numerous studies have paid attention to effects of teachers' instruction and instructional support on student academic achievement, in particular studies grounded in teacher and educational effectiveness research, and they have found clear evidence of the importance of the quality of teachers' instruction and instructional support (Muijs et al., 2014; Opdenakker, 2020). However, teacher effectiveness frameworks often recognize the importance of motivation and engagement as precursors for achievement. Therefore, it is also relevant to see whether characteristics of teachers' instruction and instructional support have effects on motivational outcomes as well.

In a study of Fauth et al. (2014), which used the model of instructional quality of Klieme et al. (2009), evidence was found for the importance of cognitive activation and supportive climate (referring to teachers' constructive feedback and

encouragement as well as to teachers' warmth and friendliness) to primary school students' development of subject-related interest.

Also, studies rooted in the TTI framework and using the CLASS/CLASS-S instrument deliver information on the relevance of teacher behaviour related to instructional support. For example, Virtanen et al. (2015) demonstrated a positive relation between instructional support and student-rated and observed general behavioral engagement among lower secondary school students in Finland and Malmberg et al. (2010) also found that observed student engagement was higher in lessons with high instructional support. However, surprisingly, Pöysä et al. (2019), investigating relations between observed instructional support in relation to a variety of situation-specific engagement indicators in Finnish grade-7 mathematics and language art classes, did not find a significant effect of (class-level) instructional support on situation-specific engagement.

Based on self-determination theory and using the TASC (student perceptions), Lietaert et al. (2015), Opdenakker (2021), and Opdenakker and Maulana (2010) found evidence for positive effects of students' perceptions of structure support on (growth in) academic engagement in the seventh grade (first year in secondary education) in Flanders (Belgium) and the Netherlands. Also, research of Hospel and Galand (2016), investigating effects of structure (and autonomy support) on behavioral, emotional, and cognitive engagement in secondary education in Belgium (French-speaking part), demonstrated clear positive associations with students' engagement (all aspects). In addition, Skinner and Belmont (1993), studying relations between student perceptions of structure, autonomy support and involvement and behavioral engagement in primary education, found evidence for the importance of (unique) effects of structure, and Opdenakker and Minnaert (2011, 2014) found, respectively, positive effects of the teacher as a helpful and good instructor and of students' basic need fulfilment of competence by the teacher on primary school students' engagement. Also, the study of Lazarides and Rubach (2017) in secondary school classes in Berlin (Germany) showed that support for competence predicted intrinsic motivation and effort (via students' mastery goal orientation). Maulana et al. (2016) found positive effects of clarity of instruction on students' intrinsic value for the subject and self-efficacy and negative effects on test anxiety in secondary education in the Netherlands. Also, Opdenakker (2013, 2014) and Stroet et al. (2015), investigating student motivation and academic engagement in prevocational and general secondary education in the Netherlands, found evidence for the importance of structure.

In addition, the study of Opdenakker (2021) revealed negative effects of chaos and inconsistency, which is often seen as the opposite of structure, on students' engagement. Furthermore, her study revealed evidence for differential effects of structure (but not of chaos/inconsistency) indicating that boys were more sensitive to structure than girls in relation to their engagement. However, the study of Lietaert et al. (2015) did not reveal evidence for this. Furthermore, research of Opdenakker and Minnaert (2014) found that teachers' fulfillment of primary students' needs to feel competent, which can be realized by offering structure, was more important for initially high academic engaged students.

Intervention studies reveal that it is possible to train teachers to successfully apply the more difficult instruction and teaching activities such as adapting instruction (more) to differences between students, and, that this training also has positive effects on student outcomes. However, research also indicates that this requires focused coaching and systematic observation of teacher's teaching during 1 or 2 years (van de Grift et al., 2011).

Furthermore, a few studies addressed the topic of differential effects. For example, Opdenakker and Minnaert (2014)¹⁷ investigated differential effects of primary school teachers' fulfillment of the need to feel competent and found evidence that initially high academic engaged students are more sensitive. Other studies found differential effects of structure in secondary education mathematics and EFL classes for boys and girls in relation to engagement indicating a higher sensitivity of boys (Opdenakker, 2021). In contrast, Tucker et al. (2002) did not find gender differences in the relation between teacher structure and student engagement, nor did Lazarides and Rubach (2017) found this with regard to the relation between teachers' support for competence and student motivational outcomes.

5.4 *Learning Climate*

Next to the quality of the teacher-student(s) relationship, which makes up the relational climate in classes in addition to student-student relationships, the class learning climate is often mentioned in learning and educational effectiveness research as well in theories and research on motivation, as an important class characteristic that influences students' learning and engagement in school. Characteristics of the classroom context as well as teachers' behaviour play a role in the creation of a good learning climate, which is often defined in terms of a stimulating and safe learning climate or a study-oriented learning climate. Evidence for the effectiveness of a study-oriented learning climate in relation to motivational outcomes is found in a diversity of studies (e.g., Dumay & Dupriez, 2007); Opdenakker, 2004; Opdenakker et al., 2005; Van Landeghem et al., 2002). Also, Telli et al. (2006), using the WIHIC, found indications that task orientation, a dimension in the WIHIC that refers to the learning climate in the class, was associated with students' attitudes towards biology in Turkish secondary education. Van de Grift et al. (2017), using the ICALT, reported a clear positive relation between a safe and stimulating learning climate in teachers' secondary education classes and student engagement in these classes in South Korea and the Netherlands. Likewise, Hughes and Coplan (2018), using a composite classroom climate indicator (based on the COS-instrument) referring to the degree to which the primary school teacher is supportive and creates a positive child-centered classroom, found evidence for a positive association between

¹⁷In addition, they found differential effects of teachers' overall fulfillment of students' psychological basic needs on engagement indicating that Dutch-speaking students were more sensitive.

classroom climate and student behavioral engagement. In addition, they also found evidence for differential effects of classroom climate in relation to student gender and anxiety indicating that, in particular, boys and students with high anxious solitude were particularly susceptible to the classroom climate.

5.5 Effects of Teachers' Autonomy Support

There is clear evidence that meeting students' need to feel autonomous and teachers' autonomy support is important for students' engagement and (intrinsic or autonomous) motivation (Opdenakker & Minnaert, 2014; Ryan & Deci, 2020; Stroet et al., 2013). This evidence is clear regarding students' engagement and motivation, across multiple educational settings and cultures, and across a variety of subjects (e.g., STEM, languages, physical education). For example, Hagger et al. (2015) found evidence for the importance of teachers' autonomy support (students' perceptions) on Pakistan secondary school students' math engagement (homework completion), while the study of Tsai et al. (2008) revealed evidence for positive effects of autonomy-supportive teacher behaviour such as understanding and taking the perspectives of students (student perceptions) on students' motivation and interest in math lessons. Studies of Bieg et al. (2011) and Jungert and Koestner (2015) also found evidence of this kind of teacher behaviour in relation to intrinsic motivation in STEM subjects. Also, the studies of Black and Deci (2000), Reeve and Jang (2006), and Roth et al. (2007) revealed positive effects of autonomy support on (autonomous) motivation, while Black and Deci (2000) also found positive effects on students' perceived competence. Assor et al. (2002) found that fostering relevance (a component of autonomy support) was positively associated with student engagement. Effects of autonomy support on students' engagement and autonomous motivation were also found in numerous other studies done e.g., in Europe (e.g., Núñez & León, 2019), the US (e.g., Reeve et al., 2004; Skinner et al., 2008) and Russia (Chirkov & Ryan, 2001), and there is also some evidence of the importance of autonomy support in more advanced educational settings (see Ryan & Deci, 2020).

Also, in the Netherlands and in Flanders (Belgium) research has demonstrated positive effects of autonomy-supportive teaching behaviour on students' academic engagement in secondary education (Lietaert et al., 2015; Opdenakker & Maulana, 2010; Opdenakker, 2014, 2021) and of the stimulation of active learning¹⁸ in Dutch primary education (Opdenakker & Minnaert, 2011). The study of Hospel and Galand (2016) in the French-speaking part of Belgium, found evidence of (unique) effects of autonomy support on emotional (and behavioral) engagement; however, no significant effect on indicators of cognitive engagement were discovered.

¹⁸ It also included attention to differentiation (and was one of the dimensions of the ISTOF student questionnaire).

Research on the differential effectiveness of autonomy support in relation to student motivational outcomes is scarce. Lietaert et al. (2015) found that only boys seemed to be sensitive to autonomy support regarding their engagement in secondary education, while Opendakker (2021) found that girls seemed to be less sensitive than boys (but still significant sensitive) to autonomy support. However, Opendakker (2021) found no evidence for differential effectiveness of controlling teaching behaviour, that is often seen as the opposite of autonomy support, in relation to student gender. Regarding the stimulation of active learning and differentiation, no differential effects were found related to gender, ethnic-cultural background, and prior engagement in a study on primary school students' engagement (Opendakker & Minnaert, 2011).

In some (other) studies, effects of controlling behaviour on motivational outcomes were explored as well. In general, negative effects of controlling teacher behaviour were found on autonomous motivation (Reeve & Jang, 2006) and engagement (Opendakker, 2021). In addition, the study of Assor et al. (2005) in Israeli primary education indicated associations with motivational orientations (extrinsic motivation and amotivation), which was partially¹⁹ mediated by negative emotions (anger, anxiety, nervousness). In addition, negative effects were found on engagement. Furthermore, evidence is found that perceptions of increases in controlling teacher behaviour are related to increases in need frustration across the school year which, in turn, relate to lower autonomous motivation, greater fear of failure, contingent self-worth and avoidance of challenges (Liu et al., 2017). In addition, there is some evidence that showing disrespect (a component of autonomy thwarting) is negatively associated with students' engagement (Assor et al., 2002) and that this component has a unique effect (as well as fostering relevance) on students' engagement. There is some evidence of biological mediators at work in the effects of autonomy-supportive versus controlling teacher behaviour indicating that the exposure to a controlling teacher is associated with higher cortisol values compared to a neutral or autonomy-supportive teacher (Reeve & Tseng, 2011), while being in learning environments characterized by autonomy support and attention to relatedness is accompanied by a higher heart rate and emotional arousal indicative of greater mobilization of energy and engagement (Streb et al., 2015).

Several intervention studies indicate that it is possible to help teachers to become more autonomy-supportive, with subsequent positive student outcomes such as engagement and autonomous motivation as a result (Assor et al., 2009; Reeve et al., 2004; see also meta-analysis of Su & Reeve, 2011).

In this context, it is relevant to mention that a lot of research using the framework of SDT delivers evidence of the importance of combining autonomy support with structure (Jang et al., 2010; Vansteenkiste et al., 2012; Sierens et al., 2009; Hospel & Galand, 2016). This means that it is important for students' motivation and engagement that teachers not only consider and welcome students' perspectives, feelings and thoughts, give them choices and allow them multiple approaches and

¹⁹The mediation seemed to be stronger for girls compared to boys.

ways to do learning tasks and solve problems, but that teachers also (instructionally) support and guide their students and provide them with clear expectations, instruction(s) and constructive feedback (Jang et al., 2010; Reeve, 2009; Skinner & Belmont, 1993; Stefanou et al., 2004; Vansteenkiste et al., 2006). The combination of high teacher autonomy support and structure has been empirically associated with not only higher autonomous motivation, but also with greater use of self-regulated learning strategies and lower test anxiety, referring to respectively cognitive and emotional engagement/disengagement (e.g., Vansteenkiste et al., 2012; Sierens et al., 2009). In addition, intervention research of, among others, Kiemer et al. (2018) and Cheon et al. (2020) reveal that it is possible to train teachers to behave more autonomy and competence supportive.

5.6 Unique or Joint Effects of Teacher Behaviour Dimensions and What Matters Most in Relation to Motivational Outcomes?

Not many studies address these topics explicitly. However, when studies include several dimensions of teacher behaviour simultaneously in the model of analysis, it is possible to make inferences about the unique effects of the dimensions in relation to the investigated outcome as well as to compare the size of effects.

Overall, there is evidence for statistically significant unique effects of the distinguished teacher behaviour dimensions in instruments discussed before on motivational outcomes (e.g., Furrer & Skinner, 2003; Jang et al., 2010; Nie & Lau, 2009; Opdenakker & Maulana, 2010; Opdenakker & Minnaert, 2011; Skinner et al., 2008; Tucker et al., 2002), although clear joint effects of the dimensions are also present. The existence of joint effects is not surprising since clear associations between dimensions of teacher behaviour were already mentioned in a previous section of this chapter. Finding unique effects of teacher behaviour dimensions indicates that these dimensions operate – at least partly – independent of each other and in a unique way to students' motivational outcomes. There is also some evidence that this is the case with regard to need-supportive versus need-thwarting teacher behaviour in relation to motivational outcomes (e.g., Assor et al., 2002; Opdenakker, 2021). However, there are also a few studies that did not find unique effects for all included (positive) dimensions of teacher behaviour (e.g., the studies of Reyes et al. (2012) and Pöysä et al. (2019), using the CLASS instrument, and the study of Hospel and Galand (2016) measuring autonomy support and structure within the theoretical framework of SDT). In addition, the study of Hospel and Galand (2016) revealed that finding unique (and mutually reinforcing) effects also depends on the type of motivational outcome investigated.

This is also the case regarding the size of effects of teacher behaviour dimensions (see e.g., Skinner & Belmont, 1993), although there are some general tendencies as well. For example, there are some indications in studies investigating teachers'

instructional support or providing structure (including clarity of instruction) and classroom management/organization that the latter has smaller effects on motivational outcomes such as academic engagement and intrinsic value than providing structure, clear instruction or instructional support (Maulana et al., 2016; Opdenakker & Minnaert, 2011).

When comparing effects of emotional support (or positive teacher-student relationships or teacher involvement) with instructional support (or structure or clarity of instruction), results seem at first sight a bit mixed. For example, in some studies (e.g., Lietaert et al., 2015; Reyes et al., 2012; Stroet et al., 2015) teacher involvement is (somewhat) more important than providing structure in relation to students' engagement (or other motivational outcomes), while in other studies (e.g., Opdenakker, 2021; Opdenakker & Minnaert, 2014) the effect of providing structure is (somewhat) larger than the effect of involvement. A deeper inspection of the mentioned studies reveals that differences in student population between the studies might be an explanation, indicating that for students of lower tracks (and with more disadvantaged backgrounds) emotional support of teachers seem to be (a bit more) important than providing structure compared to students of higher tracks (and more advantaged backgrounds) in relation to motivational outcomes, although both forms of support are important for both groups. Skinner and Belmont (1993) found, according to their path analyses, that student perceptions of teacher structure were a unique predictor of students' behavioral engagement, while students' perceptions of teacher involvement were a unique predictor of students' emotional engagement. However, an inspection of the correlations revealed that differences in associations were very small, which is in line with findings of Opdenakker and Maulana (2010) in terms of explained variance by teacher involvement and structure in relation to students' (mainly behavioural) engagement during a school year and is in line with research of de Boer et al. (2016) finding the same results with regard to intrinsic motivation of gifted students in the lower grades of secondary education in the Netherlands. In addition, their study revealed that satisfying the need to feel competent was clearly the most important need to satisfy for the intrinsic motivation of these students. Furthermore, the study indicated that teacher involvement had an additional positive effect to the effect of meeting the need to feel competent on these students' intrinsic motivation.

6 Effects of Contexts and Other Antecedents on Teacher and Teaching Behaviour

Teachers do not operate in a contextual vacuum. In their classes, they are confronted with students with specific characteristics as individuals and as a group and with structural factors such as class size, they must operate in a particular school context with its own culture, climate, policies and leadership style, they have to behave in a particular educational system with its particular characteristics (e.g., mandated

curricula; student grouping system, tracking/no-tracking, etc.), educational policies, etc. In educational effectiveness research, the importance of context is recognized for several decades. For example, educational effectiveness models such as the Comprehensive Model of Educational Effectiveness of Creemers developed in the 1990s already included context factors at class, school and above, and Reynolds, a famous educational effectiveness scholar, stated in a publication in 2000 (Reynolds, 2000) that it was necessary to study the relationships between processes, outcomes, and contexts to understand how different instructional variables relate to student outcomes in different contexts. However, until now not many (educational effectiveness) studies have been conducted to identify factors operating at the context level (Kyriakides et al., 2020). This is also the case regarding relations between school level characteristics (and class level characteristics) and teacher behaviour in classes (Opdenakker, 2020). Furthermore, the studies that investigated relations between school level characteristics and learning environment/teacher behaviour did not find strong associations (Opdenakker, 2020).

A few exceptions are found in research work²⁰ on the relationship between school/classroom context/group composition and learning environment characteristics (including teacher behaviour) (e.g., of Battistich et al., 1995; Crosnoe & Johnson, 2011; Johnson & Stevens, 2006; Maulana et al., 2016; Opdenakker, 2004; Opdenakker & Van Damme, 2006). In general, indications are found that classes and schools with favorable student populations (with regard to cognitive ability, SES, parental involvement or ethnical background) often have more favorable learning environments including more instructional support (see e.g., Opdenakker, 2004, 2019; Opdenakker et al., 2005; Opdenakker & Van Damme, 2006), more clarity of instruction (e.g., Maulana et al., 2016; Opdenakker, 2019), and a more favorable relational climate in the class (including the relationship between teacher and students and peer relations) (Opdenakker, 2004; Opdenakker & Van Damme, 2006). There is also some evidence of a less decrease in autonomy support during the school year in classes with a favorable student (ability) composition compared to classes with a less favorable composition (Opdenakker, 2014). One of the reasons could be that less favorable student populations are more challenging because they are less inclined to cooperate with teachers.

In addition, also individual student characteristics seem to matter. For example, research of Skinner and Belmont (1993) revealed a positive relationship between signs of students' engagement and the likeliness that their teachers are involved and display greater autonomy support, and more structure (contingency and consistency). Teachers respond to students who are more passive with correspondingly more neglect, coercion, and even inconsistency. When students seem to be disengaged, their teachers are less likely to provide need-supportive teaching (Escrivá-Boulley et al., 2021), exhibit more control and less autonomy support over time (Jang et al., 2016). Connell and Wellborn (1991) mentioned that teachers reported

²⁰An overview of this research with regard to Flanders (Belgium) and the Netherlands of the last three decades can be found in Opdenakker (2020).

themselves that they were less involved and offered less autonomy support to disaffected students.

Furthermore, school factors such as cooperation between teachers, school leadership style, constraints at work (e.g., accountability policy), and student-teacher ratio seem important. For example, research of Opdenakker and Van Damme (2006, 2007) revealed that cooperation between teachers at school is positively related to the quality of the relational and learning climate in classes (including teacher-student relationships), and that the school leader leadership style (namely the degree to which the leader uses a participative style and is professionalism-oriented with regard to the teachers) seems to be of importance for teachers' instructional support to their classes. In addition, evidence is found for a negative relation between constraints at work (e.g., experiencing a pressuring school environment) and teachers' psychologically controlled teaching behaviour (Soenens et al., 2012). In the same vein, research of Deci et al. (1982) has shown that the use of controlling teaching practices increases when teachers are under pressure (for example, when teachers are evaluated on students' achievement level), indicating that school systems using frequent comparative achievement tests might be pushing their teachers to rely on directly controlling teaching practices. Also, research of Pelletier et al. (2002) indicates that pressures from above (e.g., when teachers must comply with a curriculum, with colleagues, and with performance standards) is associated with more controlling and less autonomy-supportive teacher behaviour because teachers become less self-determined toward teaching. Furthermore, Ryan and Deci (2020) mention negative effects of an excessive emphasis on grades, performance goals, and pressures from high-stakes tests on teachers (and students). In addition, Cipriano et al. (2019) found that student-teacher ratio at school level was negatively associated with student perceptions of teacher support. Furthermore, research of Escrivá-Boulley et al. (2021) indicated that need-thwarting teacher behaviour was positively predicted by pressure to display authority and beliefs about the effectiveness of rewards, referring to a pressure at school level.

Lastly, also teacher characteristics such as teaching style, adherence to entity theory, teaching experience, teachers' motivation to teach, teachers' basic need satisfaction and teachers' job satisfaction are of importance. For example, Opdenakker and Van Damme (2006) found that a learner-centered teaching style seemed to matter regarding the amount of instructional support teachers gave to their classes as well as regarding the quality of the teacher-students relationship, and Escrivá-Boulley et al. (2021) found that teachers' adherence to entity theory predicted negatively need-supportive teacher behaviour. Cipriano et al. (2019) found positive associations between teaching experience and student perceptions of teacher support. Furthermore, research of Roth et al. (2007) revealed that teachers who were more autonomously motivated to teach were perceived by their students as more autonomy-supportive (and their students were more autonomously motivated to learn). However, Opdenakker (2019) did not find an association between teachers' motives for work and autonomy support, structure/clarity of instruction, classroom management and teacher involvement. Klassen et al. (2012) reported about studies showing that when teachers experienced more satisfaction of the need to feel related

with their students, they were more engaged and reported less emotional exhaustion. However, Opendakker (2019) did not find a relationship between feeling related or feeling autonomous and teacher behaviour, but, feeling competent and effective seemed to be positively related to classroom management. Furthermore, teachers' job satisfaction was positively related to teachers' involvement towards students.

Effects of teacher gender are seldom found (e.g., Maulana & Opendakker, 2014; Maulana et al., 2012, 2016; Opendakker, 2014; Opendakker & Maulana, 2010) and effects of subject taught are seldom studied, and if investigated, most of the time no effects are found (e.g., Maulana & Opendakker, 2014; Maulana et al., 2012; Opendakker, 2014; Opendakker & Maulana, 2010). An exception is the study of Opendakker et al. (2012) in which students in classes of female teachers perceived less proximity in their relationship with the teacher compared to students in classes with a male teacher. In addition, the study of Opendakker and Van Damme (2007) revealed that male teachers tend to maintain classroom order better than their female colleagues. In the same line, the study of Van Petegem et al. (2005) indicated that classroom leadership and friendliness were more associated with male than with female teachers. Furthermore, Opendakker (2019) found that teacher experience seems to matter only for male teachers regarding (student perceptions of) provided structure, clarity of instruction, autonomy support and teacher involvement; however, regarding classroom management, teacher experience mattered in a positive way for male and female teachers. In addition, there was evidence for differences in the average level of structure and autonomy support of math and English classes in favor of the math classes.

7 Conclusions, Reflections, Implications and Suggestions for Future Research Directions and Practice Related to Effective Teacher and Teaching Behaviour

A first finding reviewing current conceptualizations, measurements and instruments of teacher and teaching behaviour from a variety of perspectives was the number of different terms that were used to refer to classroom processes or practices and behaviour of teachers who appear to be good, successful, or effective in their teaching. A more sparing use of terms and clear definitions is preferable.

Second, the review indicated that a variety of research domains have an interest in classroom processes/practices and behaviour of teachers (and in their effects on student outcomes) and that, within these domains, instruments were developed to measure (the quality of) them. Dependent on the domain, these instruments are more/less grounded in theory; however, most of them are at least based on literature about 'what seems to work'. When comparing the instruments (and the theories on which they were grounded), there are many similarities in terms of the content of quality practices. However, there are differences regarding the number of

distinguished dimensions (sometimes named factors or domains) as well as with the names, wordings, and descriptions of the content of the dimensions leading to concepts with – to some degree – different descriptions and to different concepts with more or less the same meaning. It would be an advancement for the study of teacher behaviour and for the search for quality teaching practice if concepts were well-defined and uniformly used.

In addition, it would be a good idea to combine instruments in future research in the same study to investigate differences and similarities regarding concepts, operationalizations of concepts and effects of them on student outcomes, since this can help with further clarification and defining concepts. Furthermore, taking them together in one study also has more potential to yield a more comprehensive delineation of the phenomenon at hand. Still more work is needed regarding the conceptualization, operationalization, and the measurement of (the quality of) teaching and teacher behaviour and its dimensions. Kyriakides et al. (2020) reached a similar recommendation in their recent work on educational effectiveness research.

Third, the exploration of instruments and theories indicated that, in general, all the instruments (and theories) have in common an attention to teacher support and most of them address support in the domain of relation/emotion and the instructional domain. In most instruments and theories these are separated and in some it is conceptualized as one dimension. Based on the findings described in previous sections of this article, it is preferable to separate them not only because both measure on a conceptual level different things and (can) have different effects on (different) outcomes, but also because it is of importance to know where to work on in the context of professional development and learning.

In addition, most of the instruments/theories include a dimension (or subdimension) referring to class organization/management. Some instruments/theories also refer to other dimensions like autonomy support, cognitive activation, active learning, or attention to differences/differentiation. These dimensions are often included in the instruments to accommodate to newer understandings of learning and teaching. Since not only new theories on learning will be developed, but also learning in an online context will become more and more part of the teaching practice of teachers (due to and stimulated by the COVID-19 pandemic), it will be a challenge for researchers investigating (effects of) the behaviour of teachers and classroom processes to adapt their instruments to these new educational arrangements with corresponding teacher behaviour and teaching practice as well.

Forth, an important question addressed in one of the previous sections is if teaching (and teacher behaviour) must be considered/conceptualized as one-dimensional or as multidimensional/multifaceted. In fact, based on the findings described before, there is something to be said for both sides. Research with the ICALT instrument finds evidence for the one-dimensionality perspective, while research with other instruments often finds, although associations between the distinguished dimensions do exist, for the multidimensional/multifaceted perspective. An interesting perspective in line with the ‘more than one’ dimensionality perspective is research work on configurations (whether or not combined with the circumplex model). The

results described in the preceding sections reveal that there are, at one side, important associations between the distinguished teacher behaviour dimensions (in instruments and models) and common effects of these dimensions on motivational outcomes, and, at the other side, also evidence for unique effects (on top of the common effects) of teacher behaviour dimensions. These findings emphasize the importance of the need for more research on the dimensionality of teacher behaviour/teaching and of research on configurations and person-centered research to fully account for the importance of teachers and teaching in relation to student (motivational) outcomes.

Fifth, from the rather scarce research on the (in)stability of teaching and teacher behaviour there are indications for some instability of teaching and teacher behaviour (small to large changes) during the school year. There is evidence that, on average, the quality of teaching and teacher behaviour tends to decline from start to the end of the school year. This has implications for measuring teaching and teacher behaviour within a research context, but also within an accountability context. It is relevant to address questions like when and how many times a measurement is necessary to obtain good measurements of the quality of teaching and teacher behaviour.

Furthermore, the positive side of finding indications of some instability in teaching and teacher behaviour is that it is, at least, to some degree malleable and can be (positively) nurtured and advanced by professional development and learning and by favorable context conditions. Some work done in intervention studies, discussed in the preceding sections, underscore the malleability and potential for improvement of teaching and teacher behaviour; studies paying attention to links between teaching and teacher behaviour and context conditions also underscore this statement. Given the scarce research on the topic of (in)stability, more research is needed exploring stability and change between lessons and within teachers.

Sixth, a related question has to do with who the best informants are to obtain a good indication or description of the (quality of) teaching or the behaviour of a teacher. Findings reveal that there is not a straightforward answer on this question since it also depends on the goal of the measurement. There are indications that when this goal is to explain student outcomes, student perceptions are (most) valuable (and observatory information – if possible – can be informative as well), but when the measurement is part of a professional development and learning trajectory of teachers, a combination of teacher perceptions and student perceptions seems to be more valuable as well as a combination with observer ratings. If the study is small-scale and the objective is to get a thick description of the teaching and behaviour of a teacher in a particular context and time period, then observation information as well as student perceptions are perhaps the best option. If the objective is to measure the perspectives of all participants in a teaching and learning context and to tap different aspects of the learning environment, than measuring teacher as well as student perceptions is a good option. The implications of all this are that for future research a deliberate decision is necessary about what the objectives of the study and the measurement of teaching/teacher behaviour are in order to decide who will be the best informants on teaching and teacher behaviour.

Seventh, an exploration of research on the links between teaching and teacher behaviour and student motivational outcomes revealed that teaching and teacher behaviour matter, and that the instruments discussed in the preceding sections to tap information on teaching and teacher behaviour are valuable in this respect.

Furthermore, it became clear that, in particular, supportive teacher behaviour (emotional supportive by being involved and creating warm positive relationships with students and instructional supportive by providing structure and having clear instructive lessons) is of relevance for students' motivational outcomes. In addition, teachers' autonomy support (by which students are valued and supported to become autonomous, active and have a hand in their own learning process) is of importance as well as the creation of a positive (study-oriented) learning climate. In contrast, conflictual teacher-student relationships and neglecting or rejecting teacher behaviour as well as controlling teacher behaviour and teacher behaviour characterized by chaos and uncertainty is harmful for students' motivation and engagement.

Some studies also explored differential effectiveness issues in relation to student (background) characteristics such as gender, socioeconomical status or ethnicity. In general, some evidence has been found for the differential role of teacher (emotional and instructional/structure) support in relation to gender and motivational outcomes such as engagement, most of the time indicating that boys are more sensitive to teachers (involvement/emotional) support, provided structure, autonomy support, positive learning climate and teachers' neglective or rejective behaviour). Studies addressing differential effectiveness of teachers' (emotional) support related to racial or ethnic differences are rather scarce and results seem to be mixed, but when differences are found they seem to be in line with the academic risk hypothesis. Considering these limited (and sometimes contradictory) findings, additional research is needed to expand the knowledge base on differential effects of supportive teaching and teacher behaviour in relation to motivational outcomes.

Effects of classroom organization/management on motivational outcomes were also explored and it became clear that there is surprisingly little research on this topic. Although significant positive effects of this dimension were often found, this dimension was often not as strongly related to motivational outcomes as were the supportive dimensions of teaching and teacher behaviour. In addition, studies on differential effectiveness of this dimension were very scarce and delivered no evidence for the differential effectiveness of this dimension. For future research on the link between teaching and teacher behaviour and motivational outcomes, it seems worthwhile to explore the differential effectiveness of teaching and teacher behaviour in relation to gender. Furthermore, differential effectiveness in relation to other background characteristics, in particular from the academic risk hypothesis perspective, should be explored and perhaps a motivational risk hypothesis should be formulated.

Eight, studies investigating links between teacher behaviour, contexts and antecedents are scarce. The few studies available indicate that it is relevant to consider contextual and antecedent factors (such as student group composition and individual student characteristics, school culture, cooperation between teachers, school

leadership, constraints at work, student-teacher ratio, and teacher characteristics) in research, assessments, and debates about quality of teachers and teaching since they influence how teachers do and construct teaching. This line of thought agrees with ideas and work of Devine et al. (2013). A clear understanding of the effects of context and student (group) characteristics on teaching and teaching behaviour is needed since it is not only relevant to know what is good and effective, but also what the circumstances are under which teachers can manifest teacher behaviour that is defined as good or has proven to be effective regarding students' learning, development and particular outcomes. In addition, it is important to know when (circumstances, context, subject, or development domain) and for who (which kind of students) specific kinds of teacher behaviors or teaching styles are good and effective and to what degree. This asks for a perspective on teaching and teacher behaviour (in the classroom) that pays not only attention to teaching and teaching behaviour as being generic in nature (i.e. which can affect learning and development of all students in most contexts), but which also considers the broader context and situatedness of teaching and teachers' behaviour, and is sensitive to complex and dynamic interactions between teacher behaviour and student characteristics/behaviour, differentiated effectiveness and the dynamic nature of goodness, effectiveness and successfulness of teaching and teacher behaviour. Such a perspective has the potential to contribute to the establishment of stronger links between research on the quality and effectiveness of teachers and teacher behaviour, and the improvement of teaching and classroom practice because by considering context and student (group) characteristics, it assumes more complex relationships between teaching/teacher behaviour and student learning/development/outcomes and as such, it assumes a more realistic model of educational practice. Otherwise stated, by adapting to the specific needs of students, teachers, or student groups, it is expected that the successful implementation of effective teaching factors or teacher behaviours will increase and that this will ultimately maximize their potential effect on students' learning, behaviour, learning outcomes, and development.

In addition, such a perspective has the potential to help define stages of effective teaching and teacher behaviour in relation to (a diversity of) realistic educational settings and links it with equity issues as well since it takes into account differential effectiveness in relation to student (group) characteristics. The dynamic model of educational effectiveness of Creemers and Kyriakides (2008) can be seen as one of the first attempts to develop such a perspective in relation to teacher effectiveness. However, more research and theoretical work is needed to elaborate on the mentioned perspective in relation to (dimensions, dimensionality, and stages of) teaching and teacher behaviour in a diversity of educational settings (including educational levels and stages of schooling) and regarding a diversity of student outcomes and development. This will offer a more fine-grained conceptualization of effective teaching and teacher behaviour, and a more fine-grained insight in the (differential) effectiveness and successfulness of teaching and teacher behaviour, and in the underlying mechanisms and the conditions under which they can operate and contribute to equity in education. Such a perspective has the potential to address the

complex nature of (effective) teaching in a more realistic way compared to most current perspectives. In addition to theoretical work, research is needed to investigate effects of characteristics and circumstances of above school level contexts such as educational systems on teaching and teacher behaviour. To realize this, international studies are also needed.

The literature reviewed in the preceding sections gives an overview of current conceptualizations, theories, operationalizations, instruments and research addressing (the quality of) teaching and teacher behaviour and provides clear evidence of the importance of teaching and teacher behaviour in relation to (the development) of student motivational outcomes such as autonomous and intrinsic motivation and student engagement. Teachers' emotional support, involvement, quality of relationship with students, instruction, provision of structure/instructional support, the learning climate they create in their classes, their autonomy support and, to a lesser extent, also their classroom management and organization are key features accounting for links with students' motivational outcomes. In addition, evidence is delivered that teachers seem to matter even more for specific students (such as boys and vulnerable students). Positive is the finding from intervention studies that teachers can be trained to become better and more supportive teachers. Together these findings endorse the importance of investing in teacher education and teacher professionalization and to focus on the just mentioned teacher and teaching behaviour dimensions since they can stimulate students' (development of) autonomous and intrinsic motivation and engagement for school, which are important for students' achievements in school and later life. The discussed instruments to measure teacher and teaching behaviour can be helpful tools to get an idea of current practices of teachers and to have a starting point for discussions about current and future practice with and between (student) teachers.

There is from a research point of view, however, still a lot of work to do and much about teachers' significance (in a positive and a negative way) towards the development of students' motivation and engagement is not well-understood yet. Continued efforts are needed to integrate findings and research from the variety of domains discussed above to produce new research and new research findings that can help to further our understanding of development processes related to motivation and engagement (and other student outcomes) and of ways in which teachers can help (and can be helped) to ameliorate, facilitate and avoid the hindering of these developments. In addition, the use of more holistic approaches to the study of teaching and teacher behaviour (e.g., the search for configurations) is important as well as the adoption of experimental designs within real classroom settings to study and test (normative) configurations of teaching, teaching strategies and (the improvement of) teacher behaviour. Lastly, it is essential to remember that what happens in classrooms is dependent upon complex interactions between teachers and students, each with its own individual characteristics, the context they are in, and time. This implies the use of more complex models such as cross-lagged panel and dynamic longitudinal designs in future research and further theory development as well.

Appendix

Appendix Instruments Tapping Teacher Behaviour

Classroom Assessment Scoring System (CLASS)

Observation instrument based on the Teaching Trough Interactions Framework (Hafen et al., 2015; Hamre et al., 2013) and originally validated in the USA (variants for pre-K, primary and secondary education). Nowadays widely used and validated in a diversity of cultural contexts outside the USA (except for the latest version for secondary education) such as South America (Leyva et al., 2015) and Europe (Pakarinen et al., 2010).

Focus is on the patterns of interactions between teachers and students in class (because they are seen as central drivers for student learning). Support in and organization of classrooms is scored, but reference is made to teachers' behaviour related to three domains.

Emotional support: the existence of warm and caring relationships between teacher and students and enjoyment and emotional connections between teacher and students, and among students (positive classroom climate); availability of a responsive teacher who has regard for student perspectives and is sensitive to and tries to meet students' academic, affective, and social needs, who helps students resolve problems and who supports positive relations between students. A highly emotional supportive teacher has warm emotional connections with students and cares for them and consistently encourages students, provides comfort and reassurance and acts while considering their interest, motivation, and points of view.

Classroom organization: routines and procedures related to the organization of the classroom and the management of students' behaviour, time, and attention during classroom time. High scores refer to the existence of consistent schedules, established routines, a well-organized classroom, appropriate guidance, and the creation of a learning environment that is characterized by stability, predictability, and supportiveness for learning.

Instructional support: teacher's actions to support students' learning and engagement and to maximize their learning opportunities. It entails the way in which the teacher implements the curriculum to promote cognitive development, makes concepts and skills relevant to students' lives, encourages students to learn by asking questions and providing students with appropriate help and feedback that acknowledges their students' effort. Teacher activities to help students understand the content and the stimulation of higher order thinking and the delevrage of opportunities to applicate knowledge in novel contexts are included as well.

What Is Happening In this Class (WIHIC)

Student perception questionnaire (Fraser et al., 1996) (56 items) with roots in learning environments research; combines salient scales from existing questionnaires (available in the nineties) with new dimensions which became relevant at the end of the nineties; measures seven dimensions including student involvement. Four dimensions refer to a caring learning environment namely student cohesiveness, teacher support, cooperation, and equity. The other dimensions are investigation and task orientation. The original questionnaire was constructed and validated in Australia, but the final version was validated in a variety of other countries (e.g., Greece, Australia; Turkey; Asian countries e.g., Taiwan, Brunei, Singapore, Korea, China; Jordan; South-Africa; Myanmar, India, UAE) and was used for international comparisons of science classes. In contrast to other instruments discussed in this review, not all the items (and dimensions) are formulated in terms of teacher behaviour.

Student cohesiveness: the extent to which students know each other and have positive and supportive relationships with each another.

Teacher support: taking a personal interest in students (and their feelings), befriending and helping them when they have trouble with schoolwork.

Cooperation: extent to which students cooperate with each other (e.g., on assignments) during class activities.

Equity: equal treatment by the teacher regarding encouragement, help, and opportunities to be included in discussions.

Task orientation: students' attitudes towards the completion of planned activities and staying on the subject matter (including importance to get a certain amount of work done or to understand class work) and knowing the class goals.

Involvement: students' attentive interest and participation in class (e.g., giving opinions during class discussions, asking questions)" and teachers' activation of students' involvement (by asking questions or asking to explain things).

Investigation: extent to which there is emphasis on skills of inquiry and if they are used in problem solving and investigation.

International Comparative Analysis of Learning and Teaching (ICALT) Instrument

Observation instrument originally developed in and for an international context to investigate the quality of teaching (van de Grift, 2007; Maulana et al., 2021) by members of the inspectorate of the Netherlands, Belgium (Flanders), England and Germany (Lower Saxony); based on mainly earlier reviews of educational/teacher effectiveness research and existing observation instruments teaching quality evaluation. Although originally developed for evaluation purposes and inspectors' use during classroom visits in primary education, it is valid to use in secondary education (and in a variety of other countries, see Maulana et al., 2021; van de Grift, 2014;

van de Grift et al., 2017) as well, as recent research reveals (e.g., Maulana et al., 2017).

The high-inference event sampling instrument consists of 32 high-inference observable teaching acts belonging to six domains of teaching behaviour and are accompanied with 120 low-inference observable teaching activities which are considered as examples of good practices associated with the corresponding high-inference teaching act. The original ICALT distinguishes between five observable domains²¹ (with standards and corresponding indicators of good and effective teaching), namely efficient safe and stimulating learning climate, efficient classroom management, clear instruction, teaching learning strategies and adaptive teaching (adapting instruction and assignments) (van de Grift, 2007). In the adapted version (see e.g., van de Grift et al., 2014), a sixth dimension, namely activating teaching was added.

Safe and stimulating learning climate: a relaxed class atmosphere and mutual respect, and an orderly climate and intellectually stimulating environment in which there is an achievement-oriented attitude, and the self-confidence of students is encouraged by positive teacher expectations.

Efficient classroom management: starting and finishing the lesson on time, having efficient transitions between lessons, maintaining order and efficient handling of students' misconduct, and no waste of time during the lesson.

Clear instruction/clarity of instruction: setting clear lesson objectives (and checking whether they are achieved/whether students understand the learning material), having a clear lesson structure and well-structured lessons, explaining subject matter, tools and tasks clearly, and following guidelines for direct or explicit instruction.

Teaching learning strategies: provision of temporary forms of support or scaffolds to students to help them bridging the gap between present and needed skills for achievement improvement; includes teaching cognitive and metacognitive strategies.

Adaptive teaching: adaptation of teaching to student differences (being attentive to diversity of student backgrounds and personalities) to better meet students' learning needs and to optimize the learning potential of each student, in particular weak students. Adaptation can refer to additional instruction and learning time and can be realized by using the principles of pre-teaching and re-teaching.

*Activating teaching*²²: asking questions aiming to stimulate active learning, intensive instructions and teacher behaviour aimed at the activation of students' prior knowledge and making use of 'advance organizers' (Maulana et al., 2021).

²¹Depending on the publication (e.g., van de Grift, 2007; Maulana et al., 2021) also the wordings 'categories', 'dimensions' or 'scales' are used. Opportunities to learn, monitoring pupils' results and special measures for struggling learners, were not addressed in the ICALT because they were not observable in (almost) each lesson and/or most important decisions were taken at school level.

²²In the original version, this belonged to the domain 'clear instruction' (see e.g., van de Grift, 2007), which is renamed as 'clarity of instruction' in more recent publications (see e.g., van de Grift et al., 2014; Maulana et al., 2021).

The International System for Teacher Observation and Feedback (ISTOF) Instruments

Originally an observation instrument developed by an international team (and country teams) of 20 participating countries (with at least some representation of regions including North and South America, Europe, East Asia, South Asia, Southeast Asia, and Africa) during the International System for Teacher Observation and Feedback (ISTOF) project (Teddle et al., 2006).²³ In the development phase, an iterative Delphi technique drawing on expert opinion and review was used to ensure cross-cultural relevance and validity (Muijs et al., 2018). Later, the ISTOF instrument has been validated and used in other country settings as well (see for a discussion, Lindorff et al., 2020; Muijs et al., 2018).

The ISTOF instrument draws on teacher/educational effectiveness research evidence and frameworks and expert opinion and is aimed at measuring teacher effectiveness in a reliable and valid way in an international context and providing opportunities for cross-country comparisons as well as possibilities for providing meaningful feedback to teachers (Teddle et al., 2006; Kyriakides et al., 2020). The final observation instrument consists of seven (observable) components with for each component two to four indicators and for each indicator two items (45 high-inference items in total). The validity and reliability of the instrument were successfully established in a range of different contexts internationally (Muijs et al., 2018). However, in some studies the seven-components structure was not found indicating that the structure seems to be to some degree subject to variation across studies. and in some studies evidence was found for an overarching higher-order effectiveness factor as well (for a discussion, see Muijs et al., 2018).

The seven components are classroom climate, classroom management, clarity of instruction, instructional skills, promoting active learning and developing metacognitive skills, differentiation and inclusion, and assessment and evaluation. The first two belong to the overarching/super-component classroom environment, the next four ones to quality of teaching, and the last two to adaptive teaching (Teddle et al., 2006).

Classroom climate: classroom environment created by the teacher in which all students are valued, the teacher interacts with all students, communicates high expectations and initiates active interaction and participation of the students.

Classroom management: teachers' effective dealing with misbehaviour and disruption, maximization of learning time and clarity of rules.

Clarity of instruction: well-structured lessons, clear explanation of the lesson purpose, clear communication and regularly checking for understanding by the teacher.

Instructional skills: teacher's ability to engage students, possession of good questioning skills and use of various teaching methods and strategies.

²³In their article as well as in the article of Muijs et al. (2018), a detailed discussion can be found on how the ISTOF instrument was developed.

Promoting active learning and developing metacognitive skills: teacher's help to students to develop problem-solving and metacognitive strategies, giving students opportunities to be active learners, fostering critical thinking and connecting course material to students' real-world experiences.

Differentiation and inclusion: taking full account of student differences (e.g., by offering additional opportunities for practice for students who need them or by differentiating regarding the scope of assignments) and creating an environment in which all students are involved.

Assessment and evaluation: degree to which the assessment is aligned with goals and objectives and the teacher gives explicit, detailed, and constructive feedback.

In general, the ISTOF observation instrument contains components referring to more traditional approaches to teaching and learning as well as to more recent approaches. For example, classroom climate, classroom management and clarity of instruction are explicitly related to established teacher effectiveness models and research supporting direct or explicit instruction, while the components promoting active learning and metacognition, and differentiation have a link to constructivist approaches which underscore the importance of self-regulated learning (Muijs et al., 2018); the component instructional skills entail elements of both traditions.

In addition to and in close alignment with the observation instrument, Van Damme and Opdenakker developed for Flanders (Belgium) a student questionnaire (Opdenakker, 2020). This questionnaire was slightly adapted for use in the Netherlands as well (see, Opdenakker & Minnaert, 2011). The student questionnaire (46 items) revealed to have a three-factor structure and the quality of the instrument regarding the reliability of the scale scores was good. The three factors are the teacher as a helpful and good instructor (having good instructional skills, offering help and clear instruction), the teacher as promoter of active learning and differentiation, and the teacher as manager and organizer of classroom activities. Examples of items are for the *teacher as a helpful and good instructor*, 'When students encounter difficulties with the subject matter, they get help and are told what they can do to overcome these difficulties,' 'The lessons are well structured and organized,' and 'The instruction is clear and understandable.' Examples of items for the *teacher as promoter of active learning and differentiation* are, 'Examples given by students are used during class,' 'We are invited to give our personal opinions on certain subjects,' and 'Our class is divided into different groups according to the tasks given to the students.' Examples of items referring to the qualities of the *teacher as manager and organizer of classroom activities* are, 'Our classroom is often out of control' (reverse scored), and 'Most of the students are disturbed when misbehaviour occurs in our classroom.' The first mentioned factor can be interpreted as an indicator of (instructional) support and involvement of the teacher, the second one as an additional indicator of support (instructional and autonomy), and the last factor as an indicator of classroom management (Opdenakker & Minnaert, 2011).

The Teacher as a Social Context (TASC) Instruments

Questionnaires originally developed at the University of Rochester (USA) in line with the theoretical frameworks of the self-determination theory (Ryan & Deci, 2020) and the self-system process model of motivational development of Connell and Wellborn (1991). Simultaneously, a teacher and student version (for each a short and long version) were developed. Translations/adaptations and validation studies have been performed for a variety of countries (e.g., Belgium (Flanders), the Netherlands, Spain, Portugal, Indonesia) and evidence for the validity and reliability of measurements based on the TASC were reported. The long version of the student questionnaire will be addressed here (Belmont et al., 1992).

The original long-version student questionnaire consists of 52 items and taps student perceptions of teacher support and involvement referring to three dimensions: teacher involvement (14 items), structure (15 items), and autonomy support (12 items).

Teacher involvement: teacher's affection and attunement towards the student as well as teacher's dedication of resources and dependability towards the student.

Structure: teacher's help and support, adjustment and monitoring of the student, teacher's clear communication of expectations and teacher's contingency.

Autonomy support: approaching the student with respect, paying attention to the relevance of school activities and content for the student, offering choice with regard to learning and tasks and avoiding controlling behaviour and language towards the student.

References

- Aelterman, N., Vansteenkiste, M., Haerens, L., Soenens, B., Fontaine, J. R. J., & Reeve, J. (2019). Toward an integrative and fine-grained insight in motivating and demotivating teaching styles: The merits of a circumplex approach. *Educational Psychology, 111*(3), 497–521. <https://doi.org/10.1037/edu0000293>
- Aldridge, J. M., & Fraser, B. J. (2000). A cross-cultural study of classroom learning environments in Australia and Taiwan. *Learning Environments Research, 3*, 101–134. <https://doi.org/10.1023/A:1026599727439>
- Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science, 333*(6045), 1034–1037. <https://doi.org/10.1126/science.1207998>
- Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher–Student interactions in secondary school classrooms: Predicting student achievement with the Classroom Assessment Scoring System–Secondary. *School Psychology Review, 42*, 76–98. <https://doi.org/10.1080/02796015.2013.12087492>
- Anderman, L. H., & Midgley, C. (1998). Motivation and middle school students. *ERIC Digest, EDO-PS-98-5*.
- Archambault, I., Vandenbosche-Makombo, J., & Fraser, S. L. (2017). Students' oppositional behaviors and engagement in school: The differential role of the student-teacher relationship. *Journal of Child and Family Studies, 26*, 1702–1712. <https://doi.org/10.1007/s10826-017-0691-y>
- Arens, A. K., Morin, A. J. S., & Watermann, R. (2015). Relations between classroom disciplinary problems and student motivation: Achievement as a potential mediator? *Learning and Instruction, 39*, 184–193. <https://doi.org/10.1016/j.learninstruc.2015.07.001>

- Assor, A., Kaplan, H., & Roth, G. (2002). Choice is good, but relevance is excellent: Autonomy-enhancing and suppressing teacher behaviours predicting students' engagement in schoolwork. *British Journal of Educational Psychology*, 72(2), 261–278. <https://doi.org/10.1348/000709902158883>
- Assor, A., Kaplan, H., Kanat-Maymon, Y., & Roth, G. (2005). Directly controlling teacher behaviors as predictors of poor motivation and engagement in girls and boys: The role of anger and anxiety. *Learning and Instruction*, 15, 397–413. <https://doi.org/10.1016/j.learninstruc.2005.07.008>
- Assor, A., Kaplan, H., Feinberg, O., & Tal, K. (2009). Combining vision with voice: A learning and implementation structure promoting teachers' internalization of practices based on self-determination theory. *Theory and Research in Education*, 7(2), 234–243. <https://doi.org/10.1177/1477878509104328>
- Back, L. T., Polk, E., Keys, C. B., & McMahon, S. D. (2016). Classroom management, school staff relations, school climate, and academic achievement: Testing a model with urban high schools. *Learning Environment Research*, 19, 397–410. <https://doi.org/10.1007/s10984-016-9213-x>
- Battistich, V., Solomon, D., Kim, D.-I., Watson, M., & Schaps, E. (1995). Schools as communities, poverty levels of student populations, and students' attitudes, motives, and performance: A multilevel analysis. *American Educational Research Journal*, 32(3), 627–658. <https://doi.org/10.2307/1163326>
- Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1996). *Mathematics achievement in the middle school years: IEA's Third International Mathematics and Science Study (TIMSS)*. Boston College.
- Belmont, M., Skinner, E., Wellborn, J., & Connell, J. (1988). *Teacher as social context: A measure of student perceptions of teacher provision of involvement, structure, and autonomy support* (Tech. rep. no. 102). University of Rochester.
- Belmont, M., Skinner, E., Wellborn, J., & Connell, J. (1992). Teacher as social context (TASC). Two measures of teacher provision of involvement, structure, and autonomy support (*Technical Report*). University of Rochester.
- Bieg, S., Backes, S., & Mittag, W. (2011). The role of intrinsic motivation for teaching, teachers' care and autonomy support in students' self-determined motivation. *Journal for Educational Research Online*, 3(1), 122–140.
- Black, A. E., & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education*, 84(6), 740–756. [https://doi.org/10.1002/1098-237X\(200011\)](https://doi.org/10.1002/1098-237X(200011))
- Brekelmans, M., Mainhard, T., den Brok, P., & Wubbels, T. (2011). Teacher control and affiliation: Do students and teachers agree? *Journal of Classroom Interaction*, 46(1), 17–26.
- Campbell, J., Kyriakides, L., Muijs, D., & Robinson, W. (2004). *Assessing teacher effectiveness: Developing a differentiated model*. RoutledgeFalmer. <https://doi.org/10.4324/9780203403709>
- Cheon, S. H., Reeve, J., & Vansteenkiste, M. (2020). When teachers learn how to provide classroom structure in an autonomy-supportive way: Benefits to teachers and their students. *Teaching and Teacher Education*, 90, 103004. <https://doi.org/10.1016/j.tate.2019.103004>
- Chionh, Y. H., & Fraser, B. J. (2009). Classroom environment, achievement, attitudes and self-esteem in geography and mathematics in Singapore. *International Research in Geographical and Environmental Education*, 18(1), 29–44. <https://doi.org/10.1080/10382040802591530>
- Chirkov, V. I., & Ryan, R. M. (2001). Parent and teacher autonomy-support in Russian and U.S. adolescents: Common effects on Well-being and academic motivation. *Journal of Cross-Cultural Psychology*, 32(5), 618–635. <https://doi.org/10.1177/0022022101032005006>
- Cipriano, C., Barnes, T. N., Pieloch, K. A., Rivers, S. E., & Brackett, M. (2019). A multilevel approach to understanding student and teacher perceptions of classroom support during early adolescence. *Learning Environments Research*, 22, 209–228. <https://doi.org/10.1007/s10984-018-9274-0>
- Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive?* [Instructional quality: A question of perspectives?] Waxmann.

- Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In M. R. Gunnar & L. A. Sroufe (Eds.), *The Minnesota symposia on child psychology, vol. 23. Self processes and development* (pp. 43–77). Lawrence Erlbaum Associates.
- Creemers, B. (1994). *The effective classroom*. Cassell.
- Creemers, B. P. M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. Routledge. <https://doi.org/10.4324/9780203939185>
- Crosnoe, R., & Johnson, M. K. (2011). Research on adolescence in the twenty-first century. *Annual Review of Sociology, 37*, 439–460. <https://doi.org/10.1146/annurev-soc-081309-150008>
- de Boer, G., Opdenakker, M.-C., & Minnaert, A. (2016, August). *Learning environments triggering gifted students' motivation: A study on students' perception of teacher behavior, the relation with their teacher, and their motivation*. Paper presented at the international conference on motivation (ICM).
- De Jong, R., & Westerhof, K. J. (2001). The quality of student ratings of teacher behaviour. *Learning Environments Research, 4*, 51–85. <https://doi.org/10.1023/A:1011402608575>
- Deci, E.L., Spiegel, N.H., Ryan, R.M., Koestner, R., & Kauffman, M. (1982). The effects of performance standards on teaching styles: The behavior of controlling teachers. *Journal of Educational Psychology, 74*, 852–859. <https://doi.org/0022-0663/82/7406-Q8B2J00.75>.
- den Brok, P., Brekelmans, M., & Wubbels, T. (2004). Interpersonal teacher behaviour and student outcomes. *School Effectiveness and School Improvement, 15*, 407–442. <https://doi.org/10.1080/09243450512331383262>
- den Brok, P., Levy, J., Brekelmans, M., & Wubbels, T. (2005). The effect of teacher interpersonal behaviour on student subject-specific motivation. *Journal of Classroom Interaction, 40*, 20–33.
- den Brok, P., Bergen, T., & Brekelmans, M. (2006). Convergence and divergence between students' and teachers' perceptions of instructional behaviour in Dutch secondary education. In D. Fisher & M. S. Khine (Eds.), *Contemporary approaches to research on learning environments: World views* (pp. 125–160). World Scientific.
- den Brok, P., van Tartwijk, J., Wubbels, T., & Veldman, I. (2010). The differential effect of the teacher-student interpersonal relationship on student outcomes for students with different ethnic backgrounds. *British Journal of Educational Psychology, 80*, 199–221. <https://doi.org/10.1348/000709909X465632>
- den Brok, P., Mainhard, T., & Wubbels, T. (2019). Developments in quantitative methods and analyses for studying learning environments. In D. Zandvlied & B. J. Fraser (Eds.), *Thirty years of learning environments; looking back and looking forward* (pp. 41–58). Brill/Sense. https://doi.org/10.1163/9789004387720_003
- Devine, D., Fahi, D., & McGillicuddy, D. (2013). What is 'good' teaching? Teacher beliefs and practices about their teaching. *Irish Educational Studies, 32*(1), 83–108. <https://doi.org/10.1080/003323315.2013.773228>
- Dorman, J. P. (2003). Cross-national validation of the what is happening in this class? (WHIC) questionnaire using confirmatory factor analysis. *Learning Environments Research, 6*, 231–245. <https://doi.org/10.1023/A:1027355123577>
- Doyle, W. (1977). Paradigms for research on teacher effectiveness. *Review of Research in Education, 5*(1), 163–198. <https://doi.org/10.3102/0091732X005001163>
- Dumay, X., & Dupriez, V. (2007). Accounting for class effect using the TIMSS 2003 eighth-grade database: Net effect of group composition, net effect of class process, and joint effect. *School Effectiveness and School Improvement, 18*, 383–408. <https://doi.org/10.1080/09243450601146371>
- Emmer, E. T., & Strough, L. (2001). Classroom management: A critical part of educational psychology, with implications for teacher education. *Educational Psychologist, 36*(2), 103–112. https://doi.org/10.1207/S15326985EP3602_5
- Escriva-Boulley, G., Haerens, L., Tessier, D., & Sarrazin, P. (2021). Antecedents of primary school teachers' need-supportive and need-thwarting styles in physical education. *European Physical Education Review, 1–20*, 961–980. <https://doi.org/10.1177/1356336X211004627>

- Fatou, N., & Kubiszewski, V. (2018). Are perceived school climate dimensions predictive of students' engagement? *Social Psychology of Education, 21*, 427–446. <https://doi.org/10.1007/s11218-017-9422-x>
- Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction, 29*, 1–9. <https://doi.org/10.1016/j.learninstruc.2013.07.001>
- Fraser, B. J. (2012). *Classroom environment*. Routledge.
- Fraser, B. J. (2019). Milestones in the evolution of the learning environments field over the past three decades. In D. Zandvliet & B. J. Fraser (Eds.), *Thirty years of learning environments; looking back and looking forward* (pp. 1–19). Brill/Sense.
- Fraser, B. J., McRobbie, C. J., & Fisher, D. L. (1996, April). *Development, validation and use of personal and class forms of a new classroom environment instrument*. Paper presented at the annual meeting of the American Educational Research Association.
- Fraser, B. J., McLure, F. I., & Koul, R. B. (2021). Assessing classroom emotional climate in stem classrooms: Developing and validating a questionnaire. *Learning Environments Research, 24*, 1–21. <https://doi.org/10.1007/s10984-020-09316-z>
- Fuller, F. F. (1969). Concerns of teachers: A developmental conceptualization. *American Educational Research Journal, 6*, 207–226. <https://doi.org/10.3102/00028312006002207>
- Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology, 95*, 148–162. <https://doi.org/10.1037/0022-0663.95.1.148>
- Hafen, C. A., Hamre, B., Allen, J. P., Bell, C., Gitomer, D., & Pianta, R. C. (2015). Teaching through interactions in secondary school classrooms: Revisiting the factor structure and practical application of the classroom assessment scoring system-secondary. *Journal of Early Adolescence, 35*, 651–680. <https://doi.org/10.1177/0272431614537117>
- Hagger, M. S., Sultan, S., Hardcastle, S. J., & Chatzisarantis, N. L. D. (2015). Perceived autonomy support and autonomous motivation toward mathematics activities in educational and out-of-school contexts is related to mathematics homework behavior and attainment. *Contemporary Educational Psychology, 41*, 111–123. <https://doi.org/10.1016/j.cedpsych.2014.12.002>
- Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development, 72*, 625–638. <https://doi.org/10.1111/1467-8624.00301>
- Hamre, B., & Pianta, R. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development, 76*, 949–967. <https://doi.org/10.1111/j.1467-8624.2005.00889.x>
- Hamre, B. K., & Pianta, R. C. (2010). Classroom environments and developmental processes. In J. L. Meece & J. S. Eccles (Eds.), *Handbook of research on schools, schooling and human development* (pp. 25–41). Routledge.
- Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S. M., et al. (2013). Teaching through interactions: Testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *The Elementary School Journal, 113*(4), 461–487. <https://doi.org/10.1086/669616>
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Hospel, V., & Galand, B. (2016). Are both classroom autonomy support and structure equally important for students' engagement? A multilevel analysis. *Learning and Instruction, 41*, 1–10. <https://doi.org/10.1016/j.learninstruc.2015.09.001>
- Hughes, K., & Coplan, R. J. (2018). Why classroom climate matters for children high in anxious solitude: A study of differential susceptibility. *School Psychology Quarterly, 33*, 94–102. <https://doi.org/10.1037/spq0000201>
- Illeris, K. (2009). *Contemporary theories of learning: Learning theorists ... In their own words*. Routledge.

- Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *Journal of Educational Psychology, 102*, 588–600. <https://doi.org/10.1037/a0019682>
- Jang, H., Kim, E. J., & Reeve, J. (2016). Why students become more engaged or more disengaged during the semester: A self-determination theory dual-process model. *Learning and Instruction, 43*, 27–38. <https://doi.org/10.1016/J.LEARNINSTRUC.2016.01.002>
- Johnson, B., & Stevens, J. J. (2006). Student achievement and elementary teachers' perceptions of school climate. *Learning Environments Research, 9*(2), 111–122. <https://doi.org/10.1007/s10984-006-9007-7>
- Jungert, T., & Koestner, R. (2015). Science adjustment, parental and teacher autonomy support and the cognitive orientation of science students. *Educational Psychology, 35*(3), 361–376. <https://doi.org/10.1080/01443410.2013.828826>
- Kiemer, K., Gröschner, A., Kunter, M., & Seidel, T. (2018). Instructional and motivational classroom discourse and their relationship with teacher autonomy and competence support – findings from teacher professional development. *European Journal of Psychology of Education, 33*(2), 377–402. <https://doi.org/10.1007/s10212-016-0324-7>
- King, R. B. (2015). Sense of relatedness boosts engagement, achievement, and Well-being: A latent growth model study. *Contemporary Educational Psychology, 41*, 26–38. <https://doi.org/10.1016/j.cedpsych.2015.04.002>
- Klassen, R. M., Perry, N. E., & Frenzel, A. C. (2012). Teachers' relatedness with students: An underemphasized component of teachers' basic psychological needs. *Journal of Educational Psychology, 104*(1), 150–165. <https://doi.org/10.1037/a0026253>
- Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health, 74*(7), 262–273. <https://doi.org/10.1111/j.1746-1561.2004.tb08283.x>
- Klieme, E., Pauli, C., & Reusser, K. (2009). The Pythagoras study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 137–160). Waxmann.
- Konold, T., Cornell, D., Shukla, K., & Huang, F. (2017). Racial/ethnic differences in perceptions of school climate and its association with student engagement and peer aggression. *Journal of Youth and Adolescence, 46*, 1289–1303. <https://doi.org/10.1007/s10964-016-0576-1>
- Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research, 86*(3), 643–680. <https://doi.org/10.3102/0034654315626799>
- Kulik, J. A. (2001). Student ratings: Validity, utility, and controversy. *New Directions for Institutional Research, 109*, 9–25. <https://doi.org/10.1002/ir.1>
- Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research, 9*, 231–251. <https://doi.org/10.1007/s10984-006-9015-7>
- Kunter, M., Baumert, J., & Köller, O. (2007). Effective classroom management and the development of subject-related interest. *Learning and Instruction, 17*, 494–509. <https://doi.org/10.1016/j.learninstruc.2007.09.002>
- Kyriakides, L., Creemers, B., & Panayiotou, A. (2020). Developing and testing theories of educational effectiveness addressing the dynamic nature of education. In J. Hall, A. Lindorff, & P. Sammons (Eds.), *International perspectives in educational effectiveness research* (pp. 33–69). Springer. https://doi.org/10.1007/978-3-030-44810-3_3
- Lam, S.-F., Jimerson, S., Kikas, E., Cefai, C., Veiga, F. H., Nelson, B., et al. (2012). Do girls and boys perceive themselves as equally engaged in school? The results of an international study from 12 countries. *Journal of School Psychology, 50*(1), 77–94. <https://doi.org/10.1016/j.jsp.2011.07.004>

- Lazarides, R., & Rubach, C. (2017). Instructional characteristics in mathematics classrooms: Relationships to achievement goal orientation and student engagement. *Mathematics Education Research Journal*, 29, 201–217. <https://doi.org/10.1007/s13394-017-0196-4>
- Leon, J., Medina-Garrido, E., & Núñez, J. L. (2017). Teaching quality in math class: The development of a scale and the analysis of its relationship with engagement and achievement. *Frontiers in Psychology*, 8(895), 1–14. <https://doi.org/10.3389/fpsyg.2017.00895>
- Leyva, D., Weiland, C., Barata, M., Yoshikawa, H., Snow, C., & Treviño, E. (2015). Teacher–child interactions in Chile and their associations with prekindergarten outcomes. *Child Development*, 86, 781–799. <https://doi.org/10.1111/cdev.12342>
- Lietaert, S., Roorda, D., Laevers, F., Verschueren, K., & De Fraine, B. (2015). The gender gap in student engagement: The role of teachers' autonomy support, structure, and involvement. *British Journal of Educational Psychology*, 85, 498–518. <https://doi.org/10.1111/bjep.12095>
- Lindorff, A., Sammons, P., & Hall, J. (2020). International perspectives in educational effectiveness research: A historical overview. In J. Hall, A. Lindorff, & P. Sammons (Eds.), *International perspectives in educational effectiveness research* (pp. 9–31). Springer. https://doi.org/10.1007/978-3-030-44810-3_2
- Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean theorem. *Learning and Instruction*, 19, 527–537. <https://doi.org/10.1016/j.learninstruc.2008.11.001>
- Liu, J., Bartholomew, K., & Chung, P. K. (2017). Perceptions of teachers' interpersonal styles and Well-being and ill-being in secondary school physical education students: The role of need satisfaction and need frustration. *School Mental Health*, 9(4), 360–371. <https://doi.org/10.1007/s12310-017-9223-6>
- Mainhard, M. T., Brekelmans, M., den Brok, P., & Wubbels, T. (2011). The development of the classroom social climate during the first months of the school year. *Contemporary Educational Psychology*, 36, 190–200. <https://doi.org/10.1016/j.cedpsych.2010.06.002>
- Malmberg, L.-E., Hagger, H., Burn, K., Mutton, T., & Colls, H. (2010). Observed classroom quality during teacher education and two years of professional practice. *Journal of Educational Psychology*, 102, 916–932. <https://doi.org/10.1037/a0020920>
- Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective. *American Psychologist*, 52, 1187–1197. <https://doi.org/10.1037//0003-066x.52.11.1187>
- Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., et al. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, 79, 732–749. <https://doi.org/10.1111/j.1467-8624.2008.01154.x>
- Maulana, R., & Helms-Lorenz, M. (2016). Observations and student perceptions of the quality of preservice teachers' teaching behaviour: Construct representation and predictive quality. *Learning Environments Research*, 19, 335–357. <https://doi.org/10.1007/s10984-016-9215-8>
- Maulana, R., & Opendakker, M.-C. (2014). Do teacher-student interpersonal relationships deteriorate over time? An investigation of within-year changes and links with autonomous motivation in Indonesia. In D. Zandvliet, P. den Brok, T. Mainhard, & J. van Tartwijk (Eds.), *Interpersonal relationships in education: From theory to practice* (pp. 133–158). Sense. https://doi.org/10.1007/978-94-6209-701-8_9
- Maulana, R., Opendakker, M.-C., den Brok, P., & Bosker, R. (2012). Teacher-student interpersonal relationships during the first year of secondary education: A multilevel growth curve analysis. In T. Wubbels, P. den Brok, J. van Tartwijk, & J. Levy (Eds.), *Interpersonal relationships in education: An overview of contemporary research* (pp. 207–224). Sense.
- Maulana, R., Opendakker, M.-C., Stroet, K., & Bosker, R. (2013). Changes in teachers' involvement versus rejection and links with academic motivation during the first year of secondary education: A multilevel growth curve analysis. *Journal of Youth and Adolescence*, 42, 1348–1371. <https://doi.org/10.1007/s10964-013-9921-9>

- Maulana, R., Opendakker, M.-C., & Bosker, R. (2014). Teacher-student interpersonal relationships do change and affect academic motivation: A multilevel growth curve modeling. *British Journal of Educational Psychology*, *84*, 459–482. <https://doi.org/10.1111/bjep.12031>
- Maulana, R., Opendakker, M.-C., & Bosker, R. (2016). Teachers' instructional behaviors as important predictors of academic motivation: Changes and links across the school year. *Learning and Individual Differences*, *50*, 147–156. <https://doi.org/10.1016/j.lindif.2016.07.019>
- Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2017). Validating a model of effective teaching behaviour of pre-service teachers. *Teachers and Teaching*, *23*(4), 471–493. <https://doi.org/10.1080/13540602.2016.1211102>
- Maulana, R., et al. (2021). Observed teaching behaviour in secondary education across six countries: Measurement invariance and indication of cross-national variations. *School Effectiveness and School Improvement*, *32*(1), 64–95. <https://doi.org/10.1080/09243453.2020.1777170>
- McCaslin, M., & Good, T. L. (1992). Compliant cognition: The misalliance of management and instructional goals in current school reform. *Educational Researcher*, *21*(3), 4–17. <https://doi.org/10.3102/0013189X021003004>
- McLean, L., & Connor, C. M. (2015). Depressive symptoms in third-grade teachers: Relations to classroom quality and student achievement. *Child Development*, *86*, 945–954. <https://doi.org/10.1111/cdev.12344>
- Moos, R. H., & Trickett, E. J. (1974). *Classroom climate scale manual*. Consulting Psychologists Press.
- Muijs, D., & Reynolds, D. (2011). *Effective teaching. Evidence and practice*. Sage.
- Muijs, R. D., Kyriakides, L., van der Werf, G., Creemers, B. P. M., Timperley, H., & Earl, L. (2014). State of the art – Teacher effectiveness and professional learning. *School Effectiveness and School Improvement*, *25*(2), 231–256. <https://doi.org/10.1080/09243453.2014.885451>
- Muijs, D., Reynolds, D., Sammons, P., Kyriakides, L., Creemers, B. P. M., & Teddlie, C. (2018). Assessing individual lessons using a generic teacher observation instrument: How useful is the international system for teacher observation and feedback (ISTOF)? *ZDM – Mathematics Education*, *50*, 395–406. <https://doi.org/10.1007/s11858-018-0921-9>
- Nie, Y., & Lau, S. (2009). Complementary roles of care and behavioral control in classroom management: The self-determination theory perspective. *Contemporary Educational Psychology*, *34*, 185–194. <https://doi.org/10.1016/j.cedpsych.2009.03.001>
- Núñez, J. L., & León, J. (2019). Determinants of classroom engagement: A prospective test based on self-determination theory. *Teachers and Teaching*, *25*(2), 147–159. <https://doi.org/10.1080/13540602.2018.1542297>
- Opendakker, M.-C. (2004). *Leerling in wonderland? Een onderzoek naar het effect van leerling-, lesgroep-, leerkracht- en schoolkenmerken op prestaties voor wiskunde in het secundair onderwijs* [Student in wonderland? A study on the effect of student, class, teacher, and school characteristics on mathematics achievement in secondary education]. Acco.
- Opendakker, M.-C. (2013, April). *Effects on students' motivation: Does the pedagogical approach of traditional, social-constructivist and combined-approach schools matter?* Paper presented at the annual meeting of the AERA.
- Opendakker, M. C. J. L. (2014). Leerkracht-leerlingrelaties vanuit een motivationeel perspectief: het belang van betrokken en ondersteunende docenten [Teacher-student relations from a motivational perspective: The importance of involved and supportive teachers]. *Pedagogische Studien*, *91*, 332–351.
- Opendakker, M.-C. (2019, August). Quality of learning environments over time: Relationships with teacher characteristics. In J. Stang (Chair), Importance of teacher characteristics for instructional quality. Symposium conducted at the EARLI-conference.
- Opendakker, M.-C. (2020). Three decades of educational effectiveness research in Belgium and The Netherlands: Key studies, main research topics and findings. In J. Hall, A. Lindorff, & P. Sammons (Eds.), *International perspectives in educational effectiveness research* (pp. 231–286). Springer. https://doi.org/10.1007/978-3-030-44810-3_10

- Opdenakker, M.-C. (2021). Need-supportive and need-thwarting teacher behavior: Their importance to boys' and girls' academic engagement and procrastination behavior. *Frontiers in Psychology*, 12(628064), 1–17. <https://doi.org/10.3389/fpsyg.2021.628064>
- Opdenakker, M.-C., & Maulana, R. (2010, April). *Teacher-student relationships and academic engagement: How do they develop and link?* Paper presented at the international conference on interpersonal relationships in education (ICIRE).
- Opdenakker, M.-C., & Minnaert, A. (2011). Relationship between learning environment characteristics and academic engagement. *Psychological Reports*, 109, 259–284. <https://doi.org/10.2466/09.10.11.PRO.109.4.259-284>
- Opdenakker, M.-C., & Minnaert, A. (2014). Learning environment experiences in primary education: Their importance to academic engagement. In D. Zandvliet, P. den Brok, T. Mainhard, & J. van Tartwijk (Eds.), *Interpersonal relationships in education: From theory to practice* (pp. 183–194). Sense. https://doi.org/10.1007/978-94-6209-701-8_11
- Opdenakker, M.-C., & Van Damme, J. (2006). Teacher characteristics and teaching styles as effectiveness enhancing factors of classroom practice. *Teaching and Teacher Education*, 22, 1–21. <https://doi.org/10.1016/j.tate.2005.07.008>
- Opdenakker, M.-C., & Van Damme, J. (2007). Do school context, student composition and school leadership affect school practice and outcomes in secondary education? *British Educational Research Journal*, 33, 179–206. <https://doi.org/10.1080/01411920701208233>
- Opdenakker, M.-C., Van Damme, J., & Minnaert, A. (2005). *Are there equal opportunities in our classes and schools? An investigation into the relationship between class composition, indicators of the learning environment and the class climate, effort, and achievement of classes*. University of Leuven.
- Opdenakker, M.-C., Maulana, R., & den Brok, P. (2012). Teacher-student interpersonal relationships and academic motivation within one school year: Developmental changes and linkage. *School Effectiveness and School Improvement*, 23, 95–119. <https://doi.org/10.1080/09243453.2011.619198>
- Pakarinen, E., Lerkkanen, M., Poikkeus, A., Kiuru, N., Siekkinen, M., Rasku-Puttonen, H., & Nurmi, J. (2010). A validation of the classroom assessment scoring system in Finnish kindergartens. *Early Education and Development*, 21, 95–124. <https://doi.org/10.1080/10409280902858764>
- Pelletier, L. G., Séguin-Lévesque, C., & Legault, L. (2002). Pressure from above and pressure from below as determinants of teachers' motivation and teaching behaviors. *Journal of Educational Psychology*, 94(1), 186–196. <https://doi.org/10.1037/0022-0663.94.1.18>
- Pianta, R., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, 38(2), 109–119. <https://doi.org/10.3102/0013189X09332374>
- Pianta, R. C., Hamre, B. K., & Mintz, S. L. (2012). *The CLASS-Secondary manual*. University of Virginia.
- Pöysä, S., Vasalampi, K., Muotka, J., Lerkkanen, M.-K., Poikkeus, A.-M., & Nurmi, J.-E. (2019). Teacher–student interaction and lower secondary school students' situational engagement. *British Journal of Educational Psychology*, 89, 374–392. <https://doi.org/10.1111/bjep.12244>
- Reeve, J. M. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educational Psychologist*, 44(3), 159–175. <https://doi.org/10.1080/00461520903028990>
- Reeve, J. M., & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology*, 98(1), 209–218. <https://doi.org/10.1037/0022-0663.98.1.129>
- Reeve, J., & Tseng, C.-M. (2011). Cortisol reactivity to a teacher's motivating style: The biology of being controlled versus supporting autonomy. *Motivation and Emotion*, 35(1), 63–74. <https://doi.org/10.1007/s11031-011-9204-2>
- Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing students' engagement by increasing autonomy support. *Motivation & Emotion*, 28, 147–169. <https://doi.org/10.1023/b:moem.0000032312.95499.6f>

- Reeve, J., Vansteenkiste, M., Assor, A., Ahmad, I., Cheon, S. H., Jang, H., et al. (2014). The beliefs that underlie autonomy-supportive and controlling teaching: A multinational investigation. *Motivation and Emotion*, 38, 93–110. <https://doi.org/10.1007/s11031-013-9367-0>
- Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology*, 104, 700–712. <https://doi.org/10.1037/a0027268>
- Reynolds, D. (2000). School effectiveness: The international dimension. In C. Teddlie & D. Reynolds (Eds.), *International handbook of school effectiveness research* (pp. 232–256). Falmer.
- Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E., & Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality. *Learning and Instruction*, 32, 63–72. <https://doi.org/10.1016/j.learninstruc.2014.01.007>
- Roorda, D., Koomen, H., Spilt, J., & Oort, F. (2011). The influence of affective teacher–student relationships on students’ school-engagement and achievement: A meta-analytic approach. *Review of Educational Research*, 81, 493–529. <https://doi.org/10.3102/0034654311421793>
- Roth, G., Assor, A., Kanat-Maymon, Y., & Kaplan, H. (2007). Autonomous motivation for teaching: How self-determined teaching may lead to self-determined learning. *Journal of Educational Psychology*, 99, 761–774. <https://doi.org/10.1037/0022-0663.99.4.761>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and Well-being. *American Psychologist*, 55, 68–78. <https://doi.org/10.1037/0003-066x.55.1.68>
- Ryan, R. M., & Deci, E. L. (2002). An overview of self-determination theory: An organismic-dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3–33). University Press.
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. The Guilford Press. <https://doi.org/10.1521/978.14625/28806>
- Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61, 101860. <https://doi.org/10.1016/j.cedpsych.2020.101860>
- Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77, 454–499. <https://doi.org/10.3102/0034654307310317>
- Sierens, E., Vansteenkiste, M., Goossens, L., Soenens, B., & Dochy, F. (2009). The synergistic relationship of perceived autonomy support and structure in the prediction of self-regulated learning. *British Journal of Educational Psychology*, 79(1), 57–68. <https://doi.org/10.1348/000709908X304398>
- Skinner, E., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effect of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, 85, 571–581. <https://doi.org/10.1037/0022-0663.85.4.571>
- Skinner, E. A., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology*, 100, 765–781. <https://doi.org/10.1037/a0012840>
- Slavin, R. E. (2021). *Educational psychology: Theory and practice* (13th ed.). Pearson.
- Soenens, B., Sierens, E., Vansteenkiste, M., Goossens, L., & Dochy, F. (2012). Psychologically controlling teaching: Examining outcomes, antecedents, and mediators. *Journal of Educational Psychology*, 104, 108–120. <https://doi.org/10.1037/a0025742>
- Stefanou, C. R., Perencevich, K. C., DiCintio, M., & Turner, J. C. (2004). Supporting autonomy in the classroom: Ways teachers encourage student decision making and ownership. *Educational Psychologist*, 39(2), 97–110. https://doi.org/10.1207/s15326985ep3902_2
- Stipek, D. J. (1988). *Motivation to learn*. Prentice Hall.
- Streb, J., Keis, O., Lau, K., Hille, L., Spitzer, M., & Sosic-Vasic, Z. (2015). Emotional engagement in kindergarten and school children: A self-determination theory perspective. *Trends in Neuroscience and Education*, 4(4), 102–107. <https://doi.org/10.1016/j.tine.2015.11.001>

- Stroet, K., Opendakker, M.-C., & Minnaert, A. (2013). Effects of need supportive teaching on early adolescents' motivation and engagement: A review of the literature. *Educational Research Review*, 9, 65–87. <https://doi.org/10.1016/j.edurev.2012.11.003>
- Stroet, K., Opendakker, M.-C., & Minnaert, A. (2015). What motivates early adolescents for school? A longitudinal analysis of associations between observed teaching and motivation. *Contemporary Educational Psychology*, 42, 129–140. <https://doi.org/10.1016/j.cedpsych.2015.06.002>
- Su, Y.-L., & Reeve, J. (2011). A meta-analysis of the effectiveness of intervention programs designed to support autonomy. *Educational Psychology Review*, 23(1), 159–188. <https://doi.org/10.1007/s10648-010-9142-7>
- Sykes, G., & Wilson, S. (2015). *How teachers teach: Mapping the terrain of practice*. Educational Testing Service.
- Tas, Y. (2016). The contribution of perceived classroom learning environment and motivation to student engagement in science. *European Journal of Psychology of Education*, 31, 557–577. <https://doi.org/10.1007/s10212-016-0303-z>
- Tas, T., Houtveen, T., van de Grift, W., & Willemsen, M. (2018). Learning to teach: Effects of classroom observation, assignment of appropriate lesson preparation templates and stage focused feedback. *Studies in Educational Evaluation*, 58, 8–16. <https://doi.org/10.1016/j.stueduc.2018.05.005>
- Teddlie, C., Creemers, B., Kyriakides, L., Muijs, D., & Yu, F. (2006). The international system for teacher observation and feedback: Evolution of an international study of teacher effectiveness constructs. *Educational Research and Evaluation*, 22, 561–582. <https://doi.org/10.1080/13803610600874067>
- Telli, S., Cakiroglu, J., & den Brok, P. (2006). Turkish secondary education students' perceptions of their classroom learning environment and their attitude towards biology. In D. L. Fisher & M. S. Khine (Eds.), *Contemporary approaches to research on learning environments: World views* (pp. 517–542). World Scientific.
- Tsai, Y.-M., Kunter, M., Lüdtke, O., Trautwein, U., & Ryan, R. M. (2008). What makes lessons interesting? The role of situational and individual factors in three school subjects. *Journal of Educational Psychology*, 100(2), 460–472. <https://doi.org/10.1037/0022-0663.100.2.460>
- Tucker, C. M., Zayco, R. A., Herman, K. C., Reinke, W. M., Trujillo, M., Carraway, K., et al. (2002). Teacher and child variables as predictors of academic engagement among low-income African American children. *Psychology in the Schools*, 39, 477–488. <https://doi.org/10.1002/pits.10038>
- Van Damme, J., Opendakker, M.-C., & Van den Broeck, A. (2004, May). *Do classes and schools have an effect on attitudes towards mathematics?* Paper presented at the first IEA international research conference.
- van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research*, 49(2), 127–152. <https://doi.org/10.1080/00131880701369651>
- van de Grift, W. J. C. M. (2014). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement*, 25(3), 295–311. <https://doi.org/10.1080/09243453.2013.794845>
- van de Grift, W., van der Wal, M., & Torenbeek, M. (2011). Ontwikkeling in de pedagogische didactische vaardigheid van leraren in het basisonderwijs [Development in teaching skills]. *Pedagogische Studiën*, 88, 416–432.
- van de Grift, W., Helms-Lorenz, M., & Maulana, R. (2014). Teaching skills of student teachers: Calibration of an evaluation instrument and its value in predicting student academic engagement. *Studies in Educational Evaluation*, 43, 150–159. <https://doi.org/10.1016/j.stueduc.2014.09.003>
- van de Grift, W. J. C. M., Chun, S., Maulana, R., Lee, O., & Helms-Lorenz, M. (2017). Measuring teaching quality and student engagement in South Korea and The Netherlands. *School Effectiveness and School Improvement*, 28(3), 337–349. <https://doi.org/10.1080/09243453.2016.1263215>

- van der Lans, R. M., van de Grift, W. J. C. M., & van Veen, K. (2018). Developing an instrument for teacher feedback: Using the Rasch model to explore teachers' development of effective teaching strategies and behaviors. *The Journal of Experimental Education*, 86(2), 247–264. <https://doi.org/10.1080/00220973.2016.1268086>
- Van Landeghem, G., Van Damme, J., Opdenakker, M.-C., De Fraine, B., & Onghena, P. (2002). The effect of schools and classes on noncognitive outcomes. *School Effectiveness and School Improvement*, 13, 429–451. <https://doi.org/10.1076/sesi.13.4.429.10284>
- Van Petegem, K., Creemers, B. P. M., Rosseel, Y., & Aelterman, A. (2005). Relationships between teacher characteristics, interpersonal teacher behaviour and teacher wellbeing. *Journal of Classroom Interaction*, 40(2), 34–43.
- Vandenkerckhove, B., Soenens, B., Van der Kaap-Deeder, J., Brenning, K., Luyten, P., & Vansteenkiste, M. (2019). The role of weekly need-based experiences and self-criticism in predicting weekly academic (mal)adjustment. *Learning and Individual Differences*, 69, 69–83. <https://doi.org/10.1016/j.lindif.2018.11.009>
- Vansteenkiste, M., Lens, W., & Deci, E. L. (2006). Intrinsic versus extrinsic goal contents in self-determination theory: Another look at the quality of academic motivation. *Educational Psychologist*, 41(1), 19–31. https://doi.org/10.1207/s15326985Sep4101_4
- Vansteenkiste, M., Sierens, E., Goossens, L., Soenens, B., Dochy, F., Mouratidis, A., et al. (2012). Identifying configurations of perceived teacher autonomy support and structure: Associations with self-regulated learning, motivation and problem behavior. *Learning and Instruction*, 22(6), 431–439. <https://doi.org/10.1016/j.learninstruc.2012.04.002>
- Vansteenkiste, M., Ryan, R. M., & Soenens, B. (2020). Psychological need theory: Advancements, critical themes, and future directions. *Motivation and Emotion*, 44, 1–31. <https://doi.org/10.1007/s11031-019-09818-1>
- Virtanen, T. E., Lerkkanen, M.-K., Poikkeus, A.-M., & Kuorelahti, M. (2015). The relationship between classroom quality and students' engagement in secondary school. *Educational Psychology*, 35(8), 963–983. <https://doi.org/10.1080/01443410.2013.822961>
- Virtanen, T. E., Pakarinen, E., Lerkkanen, M.-K., Poikkeus, A.-M., Siekkinen, M., & Nurmi, J.-E. (2018). A validation study of classroom assessment scoring system secondary in the Finnish school context. *Journal of Early Adolescence*, 38(6), 849–880. <https://doi.org/10.1177/0272431617699944>
- Wang, M. T., & Eccles, J. S. (2012). Social support matters: Longitudinal effects of social support on three dimensions of school engagement from middle to high school. *Child Development*, 83(3), 877–895. <https://doi.org/10.1111/j.1467-8624.2012.01745.x>
- Wentzel, K. R. (2002). Are effective teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development*, 73, 287–301. <https://doi.org/10.1111/1467-8624.00406>
- Wentzel, K. R., Battle, A., Russell, S. L., & Looney, L. B. (2010). Social supports from teachers and peers as predictors of academic and social motivation. *Contemporary Educational Psychology*, 35, 193–202. <https://doi.org/10.1016/j.cedpsych.2010.03.002>
- Wolf, S. J., & Fraser, B. J. (2008). Learning environment, attitudes and achievement among middle-school science students using inquiry-based laboratory activities. *Research in Science Education*, 38, 321–341. <https://doi.org/10.1007/s11165-007-9052-y>
- Wubbels, T. (2019). My journey in the learning environments research community. Research on teacher-student interactions and relationships. In D. Zandvliet & B. J. Fraser (Eds.), *Thirty years of learning environments; looking back and looking forward* (pp. 20–40). Brill/Sense. https://doi.org/10.1163/9789004387720_002
- Wubbels, T., & Brekelmans, M. (2005). Two decades of research on teacher–student relationships in class. *International Journal of Educational Research*, 43, 6–24. <https://doi.org/10.1016/j.ijer.2006.03.003>
- You, S., & Sharkey, J. (2009). Testing a developmental–ecological model of student engagement: A multilevel latent growth curve analysis. *Educational Psychology*, 29(6), 659–684. <https://doi.org/10.1080/01443410903206815>

Marie-Christine Opdenakker is an Associate Professor/Rosalind Franklin Fellow (University of Groningen, The Netherlands) and an expert in educational/teacher effectiveness and intervention research, self-determination theory and research methodology. She studies the links between teacher behaviour/learning environment arrangements/characteristics, cognitive and non-cognitive/social-emotional student outcomes (e.g., motivation, engagement, self-regulation, procrastination, pro-/antisocial behavior) with attention to developments and differential effectiveness.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 4

Teacher Recruitment in Sweden Over the Last Two Decades: How Has Entering Teachers' GPA Changed Over Time?



Stefan Johansson

Abstract The question about what constitutes teaching quality is widely discussed in many countries and Sweden is no exception. Teaching quality has been linked to individual characteristics assumed to be related to student learning that are not necessarily associated with specialised training for the craft of teaching. One of these are the standards for entry to the profession. This chapter highlights teachers' academic performances. More specifically, it explores newly recruited teachers' grade point average over a period of over 20 years. The findings are based on register data and are analysed with descriptive statistics. The findings demonstrate how newly recruited teachers' school grade point average (GPA) has decreased the past decades but also that some quite striking differences exist depending on teachers' certification status. Implications of the results are discussed in relation to the possible effects on student achievement.

Keywords Teacher recruitment · GPA · Teacher certification · Teacher education

1 Introduction

Substantial differences in teacher effectiveness have been observed for quite some time (Darling-Hammond et al., 2005; Goldhaber & Anthony, 2007; Hanushek & Rivkin, 2012; Johansson & Myrberg, 2019; Myrberg et al., 2018; Nye et al., 2004; Rockoff, 2004). Nye et al. (2004), for example, estimated that some 10% of the variance in student achievement can be explained by the teaching quality. However, results on teacher effects are still far from conclusive and it has been claimed that teacher competence is a personal trait, little affected by education and/or that it cannot be measured by observable variables (Hanushek, 1986, 1997, 2011; Kane et al., 2008; Rivkin et al., 2005). Indeed, for some the ability to teach is a diffuse trait that cannot be predicted or particularly prepared for (e.g., Chingos & Peterson, 2011).

S. Johansson (✉)
University of Gothenburg, Gothenburg, Sweden
e-mail: stefan.johansson@gu.se

Teaching quality has been variously defined as knowing subject matter, getting high grades or test scores, being compliant and obedient, or being enthusiastic in the classroom (Darling-Hammond, 2021). These are individual qualities assumed to be related to student learning that are not necessarily associated with specialised training for the craft of teaching. However, Darling-Hammond's investigation of successful school-systems¹ around the world suggests that they do not operate on this belief. Quite the contrary, these school-systems believe that there is a distinct body of knowledge that every teacher can demonstrate and that teachers can learn to improve their performance. Darling-Hammond identified several characteristics that these school-systems had in common. One central aspect was the clear standards that outlined what teachers are expected to know and be able to do. These standards relate to the framework of teacher knowledge that Shulman (1986, 1987) described in his seminal works. Besides the standards there were other noteworthy aspects of the teacher educations of successful school-systems. For example, teacher education in these school-systems appealed to top-performing students, and attrition rates were low, both as regards entrance to the teacher education and to the profession as such. For example, in Finland entrance to preparation is highly competitive where only 10% of the applicants are admitted to preparation for primary teaching. Moreover, applicants must complete an examination that require them to read and interpret research on teaching.

In Sweden, it is quite a different situation. Teacher status has decreased in the past decades and teacher education no longer appeals to top-performing students. The declining teacher status has been under intense scrutiny by Swedish media and some years ago it was reported that student teachers were admitted to the teacher education with the lowest possible result on the SweSAT test (Örstadius, 2013). In fact, since the early 1990s has teacher students' final grades from upper secondary education declined to a significant greater extent than for other comparable groups (Alatalo et al., 2021; Bertilsson, 2014). In comparison to students in other higher education programs, student teachers have increasingly lower grades (UKÄ, 2017). One further observation is that the early dropouts from teacher education are extensive compared to other higher education programs, and it is the students with the lowest grades from upper secondary school that are dominating the dropouts (UKÄ, 2017).

At the same time as the academic achievement of the applicants decreased, there have been a number of teacher education reforms intended to raise the quality of teacher education. The many reforms that aimed for improved teacher quality have emerged during an era of expanding educational accountability including measurement and surveillance of teacher classroom behavior. While the intentions have been to raise teacher quality, the status of the profession has been on the decline, and stress and decreased job satisfaction are also increasingly observed. In order to better understand the development of teacher education in Sweden a brief background will follow.

¹Australia (with a focus on Victoria and New South Wales), Canada (with a focus on Alberta and Ontario), Finland, China (Shanghai), and Singapore.

1.1 *Teacher Education in Sweden*

In the Swedish school-system, teacher education has been reformed several times in recent decades. A nine-year long compulsory school was implemented in 1962, which resulted in a new teacher education system. Candidates opted for one of four strands aimed at, respectively, primary school grades (grades 1–3), middle school grades (grades 4–6) or towards specific subjects in secondary school grades (grades 7–9). With only minor changes, this organization lasted for some 20 years. In 1988, a new teacher education system was introduced which allowed candidates aiming to teach in compulsory schools the choice between a strand directed to primary and middle grades 1–7 and a strand directed towards the upper grades 4–9. Although the former stage-system (1–3, 4–6, or 7–9) was formally abolished in 1988, in reality the system was retained by many municipalities. As a result, teachers are not always adequately specialized for the grades they are teaching. With the beginning of 1988, and as part of a neo-liberal turn in Swedish politics, teachers were made more exchangeable and subject knowledge got an increasingly obscure position. Then again in 2001 a new teacher education reform was launched. The teacher education then went through further changes towards an increasing flexibility of teachers. The teacher education program that was launched in 2001 aimed to create a new pedagogical teacher identity where specific and well-defined subject knowledge no longer was stressed. The education became less demanding with respect to content studies and there was less emphasis on the importance of studies preparing for teaching in specific grades. A teacher could be certified to teach grades 6–12, to mention one example. In 2011, yet another teacher education reform was implemented. The pendulum had then turned towards an increasing focus on content knowledge and more specificity with respect to grade level. For example, the flexibility of the previous teacher education system with respect to teachers' subject combinations was abolished and more focused content areas were stressed (e.g., math-science combination). Teacher candidates were now to educate towards grades 1–3, 4–6 and 7–9 again. While it is challenging to quantify how the quality of teacher education has changed between the different teacher education systems, there is a possibility to shed light on the recruitment pattern to the teacher profession in Sweden using teachers' own grades. This chapter aims to describe the recruitment of teachers in Sweden during the past few decades with respect to the candidates' academic achievement. The present investigation will mainly focus on newly recruited teachers' own school grades. The research questions are:

1. How has the recruitment pattern in Swedish teacher education changed during the last two decades, with respect to teachers' average grade (GPA) levels from school?
2. How do certified and uncertified teachers' school GPA level differ?

2 Data and Method

To investigate characteristics of newly recruited teachers, data from the Swedish teacher register provided by Statistics Sweden was used. In this data, the complete population of teachers in Swedish schools is present, including detailed information about, for example, their position, their teacher education, and their certification status. In addition to the teacher register data information from The Gothenburg Educational Longitudinal Database (GOLD), which includes information about all individuals born after 1971, was added. A unique component of both registers is that it is stored by personal identification number, which facilitates a link between the teacher register and the national database GOLD, which also uses the personal identification number system. GOLD comprise rich information about individuals born after 1971, for example on their scholastic achievement. Information on GPA was added to teacher register data. Since the grading system has changed several times between 1996 and 2016, as well as grades being subject to inflation, grades were equated into percentile scores. Basically, to be in the 50th percentile means to have an average GPA in Grade 9. This study relies mainly on descriptive statistics such as mean comparisons to shed light on the general trends of teachers' grade levels over time.

3 Results

In the following the teachers' own GPA from grade 9 will be high-lighted in order to provide a picture of the recruitment pattern to the teacher profession in Sweden. Since information on GPA only is available for teachers born 1972 and later, focus is placed on specific birth cohorts or ages in the analyses. The data is cumulative in nature and more teachers are added each year. In 1996, these were just around 1500 since most teachers were older than 24. The most common ages to enter the teacher work force is 24–28 during the time-period. Some age groups were therefore selected for further analysis. In Fig. 4.1, GPA for newly recruited teachers is presented for each year, 1996–2016. To achieve comparability, different age groups (24–26 year olds and 27–29 year olds) were included.

Notably, the GPA decreases over time. A newly recruited teacher in 2010 is in the 65th percentile on average while in 1998 the same age group were in the 75th percentile. The results also suggest that teachers who join the profession earlier in life (24–26) have higher GPAs. Typically, the 24–26 year olds go from an upper secondary education to teacher education while the other age groups might have joined another profession or education before starting their teacher education. It is also worth noting that GPA mainly decreases for the teachers in the age group 24–26, and only up to about year 2005. The picture that emerges suggests that prerequisites have decreased more for those who have teaching as a first career choice. However, to also investigate the GPA by birth cohorts, those born in 1972, 1977, 1982 and

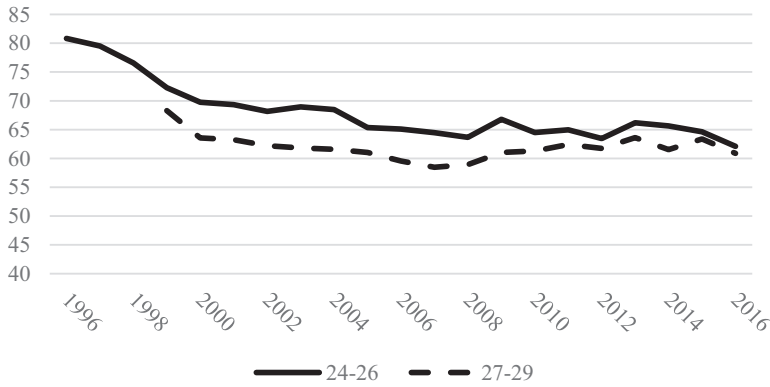


Fig. 4.1 GPA for newly recruited teachers

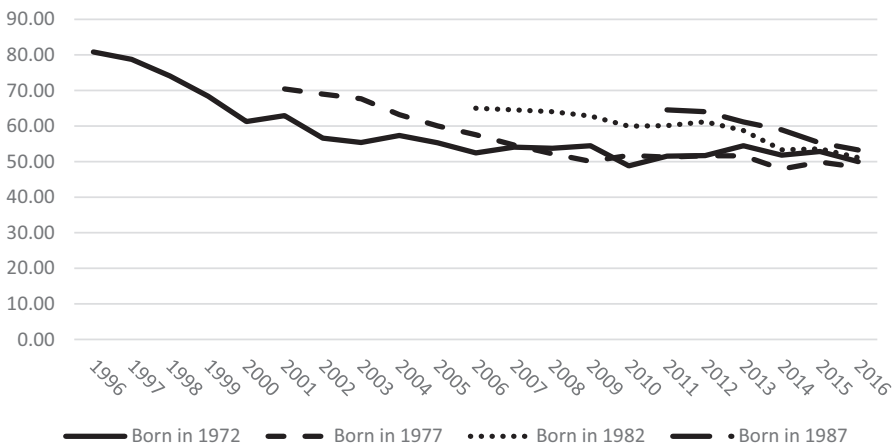


Fig. 4.2 GPA for newly recruited teachers in different birth cohorts

1987 were selected for further scrutiny. The GPA development for these cohorts is presented in Fig. 4.2.

The results demonstrate quite clearly what was seen in the previous graph; that the earlier teachers enter the profession, the better GPA they had. The pattern is quite similar for all four birth cohorts, but the older cohorts had typically better GPA in ages 24–28. When newly recruited teachers are around 30 or older the GPA lie about the 55th–50th percentile. In 2016 all teachers observed here are somewhat older than the typical entry age, some are 44 (born in 1972), others are 29 (born in 1987). However, their GPA tend to be much the same, about the 50th percentile. This is thus slightly lower than was shown in Fig. 4.1, which indicated that younger teachers were about the 60th percentile in 2016.

On the whole, the results suggest that the more able students from compulsory school have not chosen teacher education to a high extent in recent years. The

picture that emerges shows also that the recruitment to the teaching profession have gotten more homogeneous in recent past, at least in terms of grade levels. However, the grade levels are lower in the end of the period than they have been before. In the later years, there has been a large recruitment of uncertified teachers to the compulsory school, which may have led to decreases in the overall GPA. Therefore, an additional analysis to shed light on the grade development for certified and uncertified teachers respectively was conducted.

First the general trend for teachers' certification status was analysed. A large share of Swedish teachers do not hold any teacher qualifications. In Fig. 4.3 below, teachers are classified into two groups: those with a certification and those without. To be certified means that teachers have a training in education. Certification does not take into account degree of specialization and a teacher might not teach in the grades of subjects (s)he holds a training for. In the analysis of certification for two samples of teachers, the population of teachers working in grades 7–9 (Secondary) as well as teachers working in Grades 1–6 (Primary) were explored.

Figure 4.3 shows the share of uncertified teachers in the work force in secondary school (grades 7–9) and primary school (grades 1–6) respectively. This trend fluctuates somewhat across years, the general trend being that there were a higher proportion of certified teachers in the beginning of the period. In fact, the share of uncertified teachers has doubled during the time-period. It may also be noted that, teachers in primary school are certified to higher degree than is the population of teacher in secondary school 7–9. In the beginning of the 2000s the share of uncertified teachers was high, and a likely explanation of this is the large students' cohorts, and that many uncertified teachers were then hired. More teachers were hired in response to the larger student populations; however, many of these teachers did not have an adequate teacher education. The share of certified teachers has been shown to be especially low in private schools as compared with public schools. In the beginning of the 1990s, Sweden introduced a voucher system that made it more

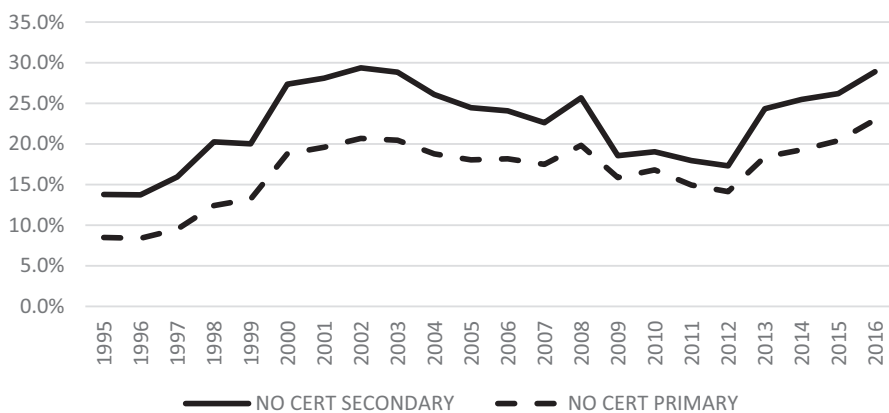


Fig. 4.3 The share of uncertified teachers in Sweden 1996–2016 divided on primary and secondary education

attractive to start new private, or independent, schools. The private schools in Sweden are tax-financed and the economic conditions are about the same as for public schools. Since the introduction of the voucher system new private schools has been introduced at an increasing rate. Much of the decision-making is delegated to school level even though more strict regulations have been formulated in recent past, for example, since 2011, a teaching license is required to assign grades. However, the teaching license have only had limited influence on the general teacher certification level. Figure 4.4 presents the share of certified teachers in public and private schools.

Based on the analyses above it could be concluded that there seems to be a need for certified teachers in Sweden. To hold a teacher training should naturally be considered as an advantage compared to have none. However, certified teachers' pre-requisites in terms of own GPA-levels need not to be higher than those of uncertified teachers. It should be noted that uncertified teachers may come from other professions that typically require higher GPA for higher-education admission than is required to enrol in teacher education. To shed light on this, an analysis of the GPA levels was carried out for certified and uncertified teachers respectively. The GPA was studied for three groups of teachers in the age of 24–26 and results are presented in Fig. 4.5. Those certified first time they teach, those who never (up to 2016) become certified, and all teachers.

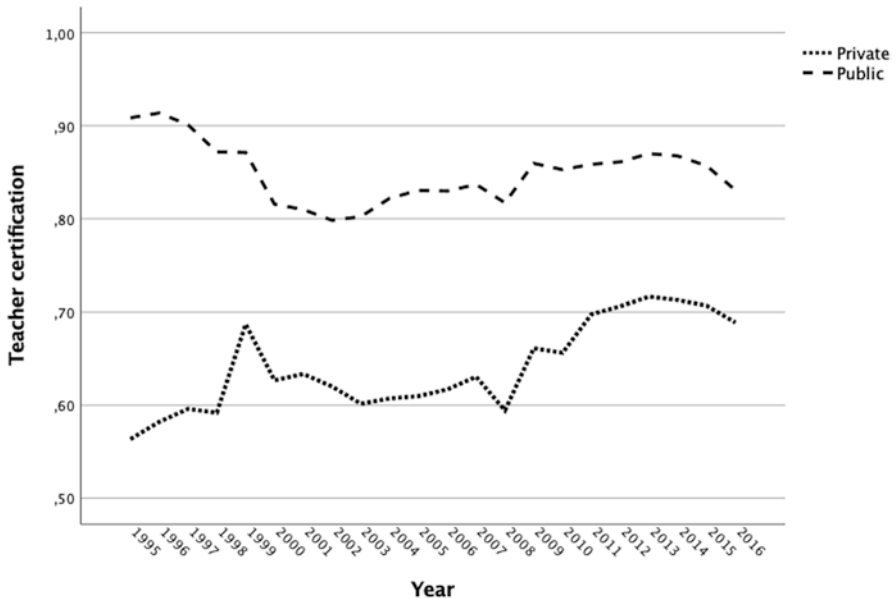


Fig. 4.4 Proportion of certified teachers in public and private schools
 As can be seen in Fig. 4.4, the proportion of certified teachers is clearly lower in the private schools. The difference is about 20% during the first decades but decreases somewhat after the teaching license requirement in 2011

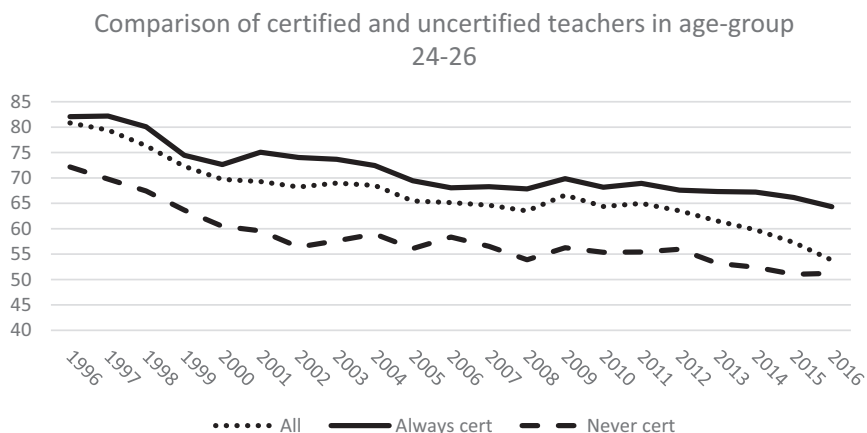


Fig. 4.5 GPA for certified and uncertified teachers

Notably, there is a general decline for all groups, thus mirroring the pattern previously shown. However, one may note that group with certified teachers has substantially higher GPA than the uncertified group. For teachers in the 24–26 years of age, who hold a certification when they start teaching, the grades decline over time but their grades are still clearly above the average (50th) percentile. Certified teachers have around the 65th–70th percentile in the last decade. The uncertified teacher group ends at about the 50th percentile. The findings are relevant to the discussion regarding the quality of certified and uncertified teachers.

It should also be noted that the group of uncertified teachers is unbalanced across time; the share is larger in the end of the period, indicated by the GPA drop in 2015–2016 for all teachers. A likely explanation for this is that many in the group of uncertified teachers are teacher candidates who did not yet receive their license but nevertheless been working in schools as teachers. The total GPA levels for teachers in Swedish compulsory school might also be affected by the entry age to the profession. In the beginning of the time-period, it was more common to start at the age of 24 and 25 than it was some years later in. One reason is due to the fact that the new teacher education 2001 was one semester longer for teachers preparing to teach grades 1–7.

4 Discussion

The picture that emerges from the register data is that the recruitment pattern of the teacher profession has changed during the past decades. Newly recruited teachers have an increasingly lower GPA from compulsory school. It is difficult to tell how

this has affected students' performance levels but a speculation is that it has contributed to the declines in Sweden's results in international comparisons.

Research has demonstrated that teacher's own schooling is important for developing both CK and PCK competencies (Kleickmann et al., 2013). Kukla-Acevedo (2009) found that only the overall GPA, not the subject specific college performance for mathematics teachers was predictive of students' 5th Grade mathematics achievement. In Swedish research, it has been difficult to demonstrate effects on students' school achievements. Grönqvist and Vlachos (2008) estimated that a decline in teachers' academic ability, expressed as aptitude test scores and final grades from upper secondary school, were negative for high-performing students, while low-performing students instead were negatively affected by having a teacher with high academic ability. While positive effects of teachers' academic ability on student achievement have been observed, international evidence is not conclusive. Harris and Sass (2011), for example, showed that elementary and middle school teachers' college entrance exam scores did not affect teacher productivity.

While it is difficult to say how the decreasing GPA levels have affected student achievement in Sweden, there are reasons to believe that the recruitment to the teacher profession has changed character with respect to the candidates' prerequisites. However, this has been a gradual change for many more years than is shown in the present study. A few studies have tried to evaluate teacher knowledge for different teacher cohorts. Alatalo (2011) used a content knowledge test, teachers' content knowledge in the Swedish language structures and basic spelling rules to examine a sample of about 300 primary-school teachers in Sweden. These teachers had substantial variation in their teacher education and years of teaching experience. The results showed that primary school teachers who qualified before 1988 (born before 1972) achieved the best test results. In another study (Frank, 2009), it was found that teachers educated before 1988 received more education in both basic and remedial reading teaching than subsequent cohorts of students, thus supporting the results of Alatalo (2011). Based on the findings of these two studies, it seems reasonable to conclude that the teachers who were educated more than 30 years ago have a more appropriate education for teaching younger pupils to read. It should also be noted that the teacher education had higher admission demands in the 1980s and that candidates likely had even higher grades than the first cohorts of the present study. However, while candidates' pre-requisites may play a role for future performances on the job, it might also interact with the quality of the teacher education and its demands.

In the present study, it was found that teachers that were somewhat older than the typical entry age (e.g., >28 years) generally have somewhat lower grades, and the share of new teachers with higher entry age has increased during the past decades. This is not necessarily negative in the sense that these teachers come with other experiences, possibly other backgrounds and different motivation. The current study used GPA from the final grade in compulsory school; however, much life experience takes place between the ages of 16 and 30, and these experiences could contribute to teachers' knowledge.

The present investigation could not relate student performances to the teachers' GPA levels. A potential drawback is that teachers from the register cannot be linked directly to their students. However, teacher and student data can be aggregated to school-level and analysed at an aggregated level. The longitudinal design allows for panel analyses where students' outcomes are measured in 3rd, 6th and 9th grade, as well as for using sophisticated multilevel models. A nice feature of the PIRLS² and TIMSS³ data is that teachers can be linked to their students; however, there is no general ability measures for the teachers in these studies. Moreover, international surveys like TEDS-m⁴ and TALIS⁵ include vast information on teachers in many countries – relating both to teacher knowledge as well as the working conditions. Both these projects are excellent in many ways but there is no link to student achievement, although successful national adaptations have been made (e.g., Baumert et al., 2010). Teacher effectiveness research is a vibrant research field and the interest in teacher quality has been intensified in the past two decades, not least with the numerous research studies accumulating. Still, however, studies including adequate controls like students' prior achievement, or studies using longitudinal and experimental designs, are rare and should be considered in future research.

References

- Alatalo, T. (2011). *Skicklig läs- och skrivundervisning i åk 1–3. Om lärares möjligheter och hinder* [Proficient teaching of reading and writing in grades 1–3. About teachers' opportunities and barriers]. Doctoral thesis, University of Gothenburg, Göteborg, Sweden. Retrieved from <http://hdl.handle.net/2077/25658>
- Alatalo, T., Hansson, Å., & Johansson, S. (2021). Teachers' academic achievement: Evidence from Swedish longitudinal register data. *European Journal of Teacher Education*, 1–21. <https://doi.org/10.1080/02619768.2021.1962281>
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and pupil Progress. *American Educational Research Journal*, 47, 133–180. <https://doi.org/10.3102/0002831209345157>
- Bertilsson, E. (2014). *Skollärare: Rekrytering till utbildning och yrke 1977–2009* [School teachers: Educational and professional recruitment 1977–2009]. Acta Universitatis Upsaliensis.
- Chingos, M. M., & Peterson, P. E. (2011). It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness. *Economics of Education Review*, 30(3), 449–465.
- Darling-Hammond, L. (2021). Defining teaching quality around the world. *European Journal of Teacher Education*, 44(3), 295–308. <https://doi.org/10.1080/02619768.2021.1919080>
- Darling-Hammond, L., Holtzman, D. J., Jin Gatlin, S., & Vasquez Heilig, J. (2005). Does teacher preparation matter? Evidence about teacher certification, teach for America, and teacher effectiveness. *Education Policy Analysis*, 13(42), 1–60. <https://doi.org/10.14507/epaa.v13n42.2005>

²Progress in International Reading Literacy Study.

³Trends in International Mathematics and Science Study.

⁴Teacher Education and Development Study in Mathematics.

⁵The Teaching and Learning International Survey.

- Frank, E. (2009). *Läsförmågan bland 9–10-åringar. Betydelsen av skolklimat, hem- och skolsamverkan, lärarkompetens och elevers hembakgrund*. [Reading skills among 9–10 year olds. The importance of school climate, collaboration between school and home, teacher competence and pupils' home background]. Acta Universitatis Gothoburgensis.
- Goldhaber, D., & Anthony, E. (2007). Can teacher quality be effectively assessed? National board certification as a signal of effective teaching. *Review of Economics and Statistics*, 89, 134–150. <https://doi.org/10.1162/rest.89.1.134>
- Grönqvist, E., & Vlachos, J. (2008). *One size fits all? The effects of teacher cognitive and non-cognitive abilities on student achievement* (IFAU Working paper 2008:25).
- Harris, D. N., & Sass, T. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics*, 95(7–8), 798–812. <https://doi.org/10.1016/j.jpubeco.2010.11.009>
- Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, 24, 1141–1177.
- Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, 19, 141–164. <https://doi.org/10.3102/01623737019002141>
- Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review*, 30, 466–479.
- Hanushek, E. A., & Rivkin, S. G. (2012). The distribution of teacher quality and implications for policy. *The Annual Review of Economics*, 4, 131–157. <https://doi.org/10.1146/annurev-economics-080511-111001>
- Johansson, S., & Myrberg, E. (2019). Teacher specialization and student perceived instructional quality: What are the relationships to student reading achievement? *Educational Assessment, Evaluation and Accountability*, 31, 1–24. <https://doi.org/10.1007/s11092-019-09297-5>
- Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, 27, 615–631. <https://doi.org/10.3386/w12155>
- Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S., & Baumert, J. (2013). Teachers' content knowledge and pedagogical content knowledge: The role of structural differences in teacher education. *Journal of Teacher Education*, 64(1), 90–106. <https://doi.org/10.1177/0022487112460398>
- Kukla-Acevedo, S. (2009). Do teacher characteristics matter? New results on the effects of teacher preparation on student achievement. *Economic of Education Review*, 28(1), 49–57.
- Myrberg, E., Johansson, S., & Rosén, M. (2018). The relation between teacher specialization and student Reading achievement. *Scandinavian Journal of Educational Research*, 63(5), 744–758. <https://doi.org/10.1080/00313831.2018.1434826>
- Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26, 237–225. <https://doi.org/10.3102/01623737026003237>
- Örstadius, K. (2013, March 27). *Nästan vem som helst kan bli lärare* [Almost anyone can become a teacher]. Dagens Nyheter. Accessed 12 Dec 2021.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73, 417–458. <https://doi.org/10.3386/w6691>
- Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 94, 247–252.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14. <https://doi.org/10.3102/0013189X015002004>
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57, 1–22. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>
- UKÄ. (2017). *Lärarstudenternas gymnasiebetyg, avhopp och studieprestation*. Statistisk analys [Teacher students' upper secondary school grades, dropouts and study performance. Statistical analysis]. Universitetsskanslersämbetet [The Swedish Higher Education Authority].

Stefan Johansson is associate professor and researcher at the University of Gothenburg. In recent research projects he has focused the meaning of teacher competence and how teacher quality can be operationalized. His previous studies have investigated the effects of different teacher competence indicators on student achievement. Furthermore, his research interests center on the use and consequences of international large-scale assessments, such as TIMSS and PISA.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 5

Effective Teaching: Linking Outcomes of Active Citizenship to Learning Environments



Gordon Sturrock and David Zandvliet

Abstract This chapter discusses the use of a learning environment instrument, the Place-Based Learning and Constructivist Environment Survey (PLACES) in an environmental studies program that operated out of British Columbia, Canada. In order to access information about students' perceptions, the instrument was implemented in an Integrated Environmental Studies program called Experiential Studies 10 (ES 10) as part of a range of evaluation methods. The study was retrospective in nature utilizing a mixed method approach to determine the long-term effects of the program on participants' citizenship activities. Our findings demonstrate that learning environment and citizenship outcomes were linked, and key learning environment features were identified as being important for long term outcomes of active citizenship. This chapter will provide a brief overview of the study and shed light on how paying close attention to the learning environment created within environmental education programming can contribute to long-term outcomes of active citizenship.

Keywords Learning environments · Active citizenship · Place based learning

1 Introduction

Contemporary learning environments research is a diverse field of inquiry and various approaches, studies and instruments have been developed, tested and validated in diverse settings and countries, with particular attention to science education

G. Sturrock (✉)

Sport Science Department, faculty of Science and Technology, Douglas College,
New Westminster, BC, Canada

e-mail: sturrockg@douglascollege.ca

D. Zandvliet

Simon Fraser University in Vancouver, Vancouver, BC, Canada

Institute for Environmental Learning, Burnaby, BC, Canada

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_5

contexts (Fraser, 1998, 2014; Zandvliet & Fraser, 2018). This research trajectory has “provided convincing evidence that the quality of the classroom environment in schools is a significant determinant of student learning” (Dorman et al., 2006, p. 2). Further, there is compelling evidence suggesting that classroom environments of various types can have a strong effect on other types of student outcomes including attitudes (Fraser & Butts, 1982; Fisher & Khine, 2006; Fraser, 2007, 2014). In this study, we explore the concept of ‘active citizenship’ as another type of outcome that is potentially influenced or predicted by the learning environment as co-constructed among teachers and students.

Today, a large amount of school time is spent in classroom environments where students are expected to learn skills to help navigate and achieve success in a global environment. Schools play a key role in shaping students to be successful in society but also prepares them to be a contributing member as an active citizen. Positive learning environments can play a large role in creating experiences that lead to long-term outcomes such as active citizenship. Active citizens can be described as people who care about their local communities and beyond. Active citizens actively embrace social responsibility and take it upon themselves to play a civic role of being informed and maintaining and developing critical perspectives while becoming actively involved in social, political and/or environmental issues (Kincheloe, 2005). Pickett and Fraser (2010) define the classroom learning environment as “the students’ and teachers’ shared perceptions” (p. 321) within the learning space created. Learning space can be described as the physical setting for learning: the place in which teaching and learning occur, which can happen indoors or outdoors. The psychosocial environment includes all relationships that exist between participants (teacher, student, and other students). The majority of research and evaluation of education includes measures of academic achievement and other learning outcomes without much reference to the educational process (Pickett & Fraser, 2010). More recently, significant progress has been made in the “conceptualization, assessment, and investigation of the learning environments of classrooms and schools” (Pickett & Fraser, 2010, p. 321). Zandvliet (2014) describes research on learning environments “as both descriptive of classroom contexts and predictive of student learning” (p. 18). Therefore, research in learning environments plays a valuable role in the field of education especially if one wants to make connections between long term outcomes. Zandvliet (2012) asserts that research in learning environments plays a valuable role in the field of education, especially the evaluation of new curricula or innovations, which would include innovative programs with citizenship outcomes. This kind of research can provide “the description of a valuable psychological and social component of students’ educational experience” (p. 18). There is convincing evidence that links the quality of the classroom environment in schools (which relates to the interpersonal interactions between the teacher and students) toward student learning, which includes achievement, attitude and behaviours (Pickett & Fraser, 2010; Zandvliet, 2014). This chapter describes a long-term study on an integrated curriculum program called *Experiential Studies 10* that demonstrates that learning environment and citizenship outcomes can be linked, and that key learning environment features can be identified as contributing to the long term outcome of

active citizenship. It begins by providing a brief overview of the study and then investigates how key learning environment features of the programs lead to long-term outcomes of active citizenship.

1.1 The Experiential Studies 10 Program

The Experiential Studies 10 (ES 10) program can be considered as an example of an integrated curriculum program. Integrated curriculum programs (ICPs) are interdisciplinary educational programs that blend content from various sources around a common theme. Typical ICPs combine various courses taught in a holistic manner. The ES 10 program is an ICP that combines Science 10, Earth Science 11, Social Studies 10, and Physical Education 10. Horwood (1994) states, “Integration happens, not so much from putting school subjects together into a shared time and space, but from certain types of general experience which transcends disciplines” (p. 91). ICPs tend to blend complementary subject areas with the intention of creating interdisciplinary investigations of a central theme, topic, or experience (Jacobs as cited in Breunig & Sharpe, 2009). The ES 10 program is an ICP that utilized a multidisciplinary and place-based education approach to foster critical thinking. The program includes a multitude of real-life learning experiences conducted in various locations in Southern British Columbia, Canada. Examples of these experiences include: working in partnership with other integrated curriculum program students, conducting various forest mapping and environmental monitoring for sustainable forest practices on Salt Spring Island and working alongside a University of British Columbia PhD candidate on a study of sea lice and salmon fry.

1.2 Place Based Education

The notion of a place-based education was described by Soble (1993, 1996) and others have expanded these ideas (Gruenewald, 2003; Hutchison, 2004; Orr, 1992, 1994; Thomashow, 1996; Woodhouse & Knapp, 2000). Describing exactly what constitutes a place-based education becomes clouded partly due to the multifaceted and interdisciplinary nature of the literature where this notion seems to reside. Gruenewald (2003) asserts that the idea of place-based learning connects theories of experiential learning, contextual learning, problem-based learning, constructivism, outdoor education, indigenous education, and environmental education. This paper relates how learning environment methodologies can be employed effectively in place-based and environmental education studies and relates the development of a valid and reliable tool for this purpose. Many benefits can be achieved by engaging students in place-based environmental education programs, these include: improvement in their academic achievement, problem solving, critical thinking, co-operative learning skills, and an increased motivation to learn (Zandvliet, 2012). In addition,

place-based practices have been demonstrated to be an important learning feature towards outcomes of active citizenship (Sturrock, 2017). Keeping this focus in view, this study reports on the use of a learning environment instrument: the Place-based and Constructivist Learning Environment Survey or PLACES (Zandvliet, 2012) as it relates to the development of students' citizenship values.

Through place-based environmental education, learners' cognitive structures may be altered, environmental attitudes modified and the general learning environment that develops around these programmes can enrich and stimulate further learning. These elements are viewed as interconnected and will change as a whole system, not as separate parts (Johnson & Johnson, 2003). This type of research has been described as congruent with an ecological view of education (Zandvliet, 2012). In this chapter, we detail a study of the students learning environment to examine how the types of learning environments developed in place-based environmental education settings as well as its association to student outcomes such as citizenship. We also consider the suitability of the PLACES instrument for environmental education research in this particular learning context.

2 Methodology

This case study uses a mixed methodology that incorporates both qualitative and quantitative research methods. The study context was a grade 10 Integrated Environmental Studies Program called Experiential Studies 10 (ES 10) from a Canadian high school. Three different cohorts from years 2003, 2004, and 2007 were included in the study. Both the 2004 and 2007 cohorts had 24 students of relatively equal number of males and females while the 2004 cohort had 23 students with 16 females and 7 males. Refer to Table 5.1 for a detailed demographic of participants from the 2003/04 cohorts. Data collection protocols included administration of quantitative surveys (PLACES), focus groups, open ended questionnaires, and participant-researcher observations. The study was also longitudinal in nature as one cohort of students were administered a learning environment survey 5 years earlier as part of an earlier study and five years later as part of a follow-up study. The first set of data collection was conducted in 2007 (Koci, 2013) and cross-referenced five years later (Sturrock, 2017). Two other cohorts from 2003 and 2004 were included in the study to provide deeper understanding of the long-term effects of program related to active citizenship. For these cohorts the PLACES survey, active citizenship survey, focus groups, and open-ended questionnaires were retrospective in nature. The core research question for this study was: "What are the perceptions of a group of alumni from a Grade 10 integrated curriculum program (ES 10) with regard to the effects of the program on their citizenship activities?". The four sub questions addressed engagement in communities or beyond, perceived influence of the program relating this engagement, skills that have been developed or fostered having a positive effect towards community participation and aspects of the program that had the greatest general impacts.

Table 5.1 Demographic of participants

Demographic of the 2003/04 ES 10 cohorts		
	2003 Cohort	2004 Cohort
Total number of students	24 students (12 male, 12 female)	23 students (7 male, 16 female)
Ethnicity of students	Majority of the students were born in Canada and were Caucasian. A small percentage (5/24) of students were immigrants, all attaining Canadian Citizenship at the time of the study. Ethnicity of immigrants: (Chinese (2), Korean (1), Russian (1), Chile (1)).	Majority of the students were born in Canada and were Caucasian. A small percentage (4/23) of students were immigrants, all attaining Canadian Citizenship at the time of the study. Ethnicity of immigrants: (Chinese (3), German (1)).
Socio-economic status of students	Most of the students came from middle class families with a small percentage (1–2 in each cohort) from the lower middle class family. Note: Majority of people living in Coquitlam, BC (in the Centennial High School area) are considered middle classed (Tri-Cities Chamber of Commerce, 2014).	
Academic profile of students	Although there was an application process for students accepted into ES 10, academic achievement was not a criterion. The students in this cohort represented an average level of academic achievement. However, since this was a unique program and students do apply to be part of it, suggests that it is self-selecting and may attract more self-motivated students. Note: Two seats were reserved for students that were identified by counsellors and/or administrators that were having difficulties in regular school usually associated with achievement, and attendance.	

To further augment the active citizenship portion of the study the International Social Survey Program (ISSP) Citizenship 2004 survey was administered to the 2003/04 cohort. The results from the ISSP Citizenship 2004 survey (ISSP, 2012) were utilized to compare values from the ES 10 group to data collected in 2004 on 47 countries, including Canada, as part as the ISSP. Comparisons include the ES 10 results compared to all ages in Canada and more importantly data from the same age group (23–24 years of age). The results from this survey indicate areas where the ES 10 group score higher or lower than the comparison groups. Since the variable list for the ISSP Citizenship 2004 survey includes constructs that can be used as indicators of active citizenship, the comparison provides an indicator of the long-term effects of the ES 10 program relating to active citizenship. These indicators include community participation, political action, empowerment, informed citizen, tolerance, and voice, which is consistent with active citizenship research (Durr, 2004).

2.1 Data Source/Evidence

The questionnaire selected for the study is one that had been tested and proven to be reliable in measuring learning environments in secondary classrooms (Zandvliet, 2012). The Place-based and Constructivist Environment Survey (PLACES) has

been extensively utilized throughout six countries and administered to over 3000 students (Zandvliet, 2007, 2012) showing consistently acceptable measures of internal consistency (Cronbach alpha reliability) and for discriminant validity for its eight constructs. Furthermore, three of the constructs from the tool (critical voice, community relevance and student cohesiveness) are significant learning environment factors that have been linked to long-term active citizenship (Ireland et al., 2006). As the questionnaire is not time or age sensitive, the questionnaire was easily adapted for our use in this study setting. The PLACES questionnaire has eight scales adapted from the previously referenced inventories and were derived from data that emerged from a qualitative study of environmental educators' preferences as such, PLACES can be described as a compendium on constructs viewed by place-based and environmental educators as being most important for their practice (Zandvliet, 2012). Table 5.2 gives sample items from each scale for the PLACES questionnaire (Zandvliet, 2012).

Data collection for our study proceeded in two phases. For the 2007 cohort, each student was asked to complete the Preferred form of PLACES within the first week of the program, and on the last day of course each student was asked to complete the Actual form of PLACES. To complete the questionnaires each statement was responded to using a Likert scale 1–5. Validity and reliability data were calculated for all samples. Five years later the original cohort was contacted again and asked to complete the Actual-PLACES questionnaire once more. Summaries of the results relating to the 2007 cohort can be found in Tables 5.3, 5.4, 5.5, and 5.6 which includes validity and reliability data. These survey results were then augmented by administering the PLACES questionnaire to the 2003 and 2004 cohorts and followed up with a group interview, individual interviews, and an open-ended questionnaire. The class size for the 2003 and 2004 cohorts were 24 and 23 respectively with 36 of these past graduates participating in the study. Refer to Table 5.7 for the summary of the PLACES results for the 2003 and 2004 cohorts. The rationale for utilizing the 2003 and 2004 cohorts was to ensure long-term results since these graduates completed the program eight to nine years earlier at the time of the data collection and that many of these students completed their post-secondary studies.

Table 5.2 Sample statements from the selected scales for PLACES questionnaire

Relevance/Integration (CI)	I want my lessons to be supported with field experiences and other field-based activities.
Critical Voice (CV)	It would be ok for me to speak up for my rights.
Student Negotiation (SN)	I want to ask other students to explain their ideas and opinions.
Group Cohesion (GC)	I want students to get along well as a group.
Student Involvement (SI)	I want to ask the instructor questions when we are learning.
Shared Control (SC)	I want to help instructors plan what I am to learn.
Open-Endedness (OE)	I want opportunities to pursue my own interests.
Environmental Interaction (EI)	I want to spend most of the time during field local trips learning about my environment.

Table 5.3 2007 Cohort pre-actual results (Perceptions of the traditional classroom)

Scale	Mean	σ	CA	DV
Relevance/Integration	2.6	0.59	0.6	0.29
Critical Voice	3.6	0.82	0.7	0.32
Negotiation	3.2	0.79	0.8	0.32
Cohesiveness	2.8	0.70	0.8	0.39
Involvement	3.2	0.73	0.7	0.27
Control	1.7	0.74	0.8	0.21
Open Endedness	3.0	0.50	0.6	0.32
Environmental Interaction	3.5	0.55	0.7	0.17

Table 5.4 2007 Cohort ES-actual results (Perceptions of the ES 10 Program)

Scale	Mean	σ	CA	DV
Relevance/Integration	4.2	0.64	0.8	0.30
Critical Voice	4.8	0.26	0.8	0.09
Negotiation	4.3	0.53	0.8	0.37
Cohesiveness	4.7	0.53	0.8	0.21
Involvement	4.2	0.50	0.6	0.41
Control	3.7	0.76	0.8	0.24
Open Endedness	4.4	0.52	0.6	0.37
Environmental Interaction	4.4	0.42	0.7	0.17

Table 5.5 2007 Cohort pre-preferred results (Preferred learning perceptions at start of ES 10)

Scale	Mean	σ	CA	DV
Relevance/Integration	4.2	0.42	0.6	0.16
Critical Voice	4.7	0.35	0.7	0.33
Negotiation	4.1	0.57	0.7	0.40
Cohesiveness	4.6	0.41	0.7	0.32
Involvement	4.1	0.60	0.7	0.35
Control	3.8	0.75	0.8	0.47
Open Endedness	4.3	0.56	0.7	0.44
Environmental Interaction	4.0	0.67	0.7	0.37

Table 5.6 2007 Cohort post results (Perceptions of program five years later)

Scale	Mean	σ	CA	DV
Relevance/Integration	4.5	0.42	0.7	0.29
Critical Voice	4.9	0.26	0.8	0.22
Negotiation	4.1	0.48	0.8	0.41
Cohesiveness	4.7	0.43	0.8	0.30
Involvement	4.3	0.49	0.7	0.34
Control	3.9	0.50	0.7	0.21
Open Endedness	4.6	0.32	0.6	0.31
Environmental Interaction	4.6	0.28	0.6	0.34

Table 5.7 2003/04 Cohorts post results (Perceptions of program eight to nine years later)

Scale	Mean	σ	CA	DV
Relevance/Integration	4.4	0.39	0.7	0.3
Critical Voice	4.7	0.33	0.7	0.36
Negotiation	4.3	0.5	0.8	0.37
Cohesiveness	4.7	0.3	0.7	0.32
Involvement	4.4	0.46	0.7	0.3
Control	3.5	0.65	0.9	0.39
Open Endedness	4.5	0.44	0.7	0.31
Environmental Interaction	4.5	0.37	0.7	0.33

The rationale for including the 2007 cohort was due to the availability of preprogram and post program data as it relates to the PLACES learning environment tool from Koci's (2013) study. The results from administering the PLACES questionnaire to the 2007 cohort five years later helps determine consistency of the instrument related to long-held perceptions (beliefs) which is significant for learning environment research and for this study since participants were asked to recall their experiences in the program that occurred eight to nine years earlier. We were able to follow up with 18 out of 24 possible students in the 2007 cohort.

3 Results

As in previous studies, the Cronbach alpha (CA) was utilized to measure internal consistency while discriminant validity (DV) was utilized to measure validity for the scales in PLACES. The Chronbach alpha calculates the internal consistency of the items within each scale or construct, which indicates that all the questions within the same construct are responded to similarly. Higher numbers represent better internal consistency with 1.0 indicating a perfect correlation. High consistency indicates the questions within the scale are responded to similarly and so can be aggregated together into one factor. Values of 0.6 or less are considered poor or unreliable (George & Mallery, 2003). The discriminant validity (DV) is used to determine if each of the eight constructs is measuring a unique (or distinct) concept. Constructs that measure something conceptually different than other scales have values of 0.4 or less (Revelle & Zinbarg, 2009). The calculated values from the Cronbach alpha and discriminant validity data from administration of PLACES across the time frame of this study indicated that that the eight constructs included in both forms of the instrument demonstrated acceptable within scale reliabilities but also discriminated validly among the eight constructs measured. This demonstrates that the PLACES instrument is robust and was suitable for use within the context of our study. Tables 5.3, 5.4, 5.5, and 5.6 highlight students' perceptions for the 2007

cohort as described by the PLACES instrument at various times over the course of this longitudinal study which also includes Cronbach alpha and discriminant validity data (all within the acceptable range as described above).

In each setting, the mean responses for each scale of the preferred questionnaire (Table 5.5) are similar to the responses for the actual form of the questionnaire (Table 5.4), thereby confirming the findings of our preliminary case study work. This indicates that students' actual learning environment often met the expectations of their preferred learning environment as measured by the PLACES questionnaire. Overall, these data indicates that students were more satisfied with the learning environments created through the experiential programmes than they were with the learning environments created through more traditional classroom-based programmes.

In general, study results also describe how student participation in this type of programme might change students' expectations for overall learning and for the educational learning environments they encounter in schools and provide rich (more holistic) descriptions of the different learning environments experienced by students. Another key finding was that students' perceptions were very stable over the long timeframe of this study (5 years) and that certain aspects of the learning environment were closely associated with Citizenship outcomes. Table 5.6 demonstrates the PLACES results five years later while Fig. 5.1 displays the ES 10 participants perception results in a graph format five years later to the actual program results. The two graphs are remarkably similar demonstrating how stable student's perceptions using the PLACES inventory was over a five-year time period.

The PLACES survey tool was also utilized for the ES 102003/04 cohorts to assess students' perceptions of their learning environment while in ES 10, administered eight to nine years after being in the program. The PLACES results for the 2003/04 cohorts are shown in Table 5.7 which also includes Cronbach alpha and discriminant validity values (all in the acceptable range). The information from the PLACES survey indicated learning environment features that students feel are important that lead to long-term learning and active citizenship. The overall mean score (sum mean of all data) for the 2003/04 cohort was 4.4, indicating a positive perception of the ES 10 learning environment by the graduates of this program. Comparing the 2007 cohort results from Koci's (2013) study to the same group of students five years later (2007 cohort post 5 years) shows striking similarity in values. The overall mean score for the 2007 cohort from Koci's (2013) study was 4.4 while the overall mean score from the same group of students five years later was 4.5.

The qualitative portion of this study included a focus group and individual interviews for participants not available for the group interview, and an additional open-ended questionnaire. The focus group method utilized an Interview Matrix method (Chartier, 2002). The 2003 and 2004 ES 10 cohorts formed a large focus group of 21 students. The interview matrix is a tool to build dialogue for groups of up to 40 participants. The methodology allows for full engagement in dialogue, equal participation, focused discussion and consensus building. Both cohorts were interviewed at the same time to help limit recall effects associated with a single "familiar"

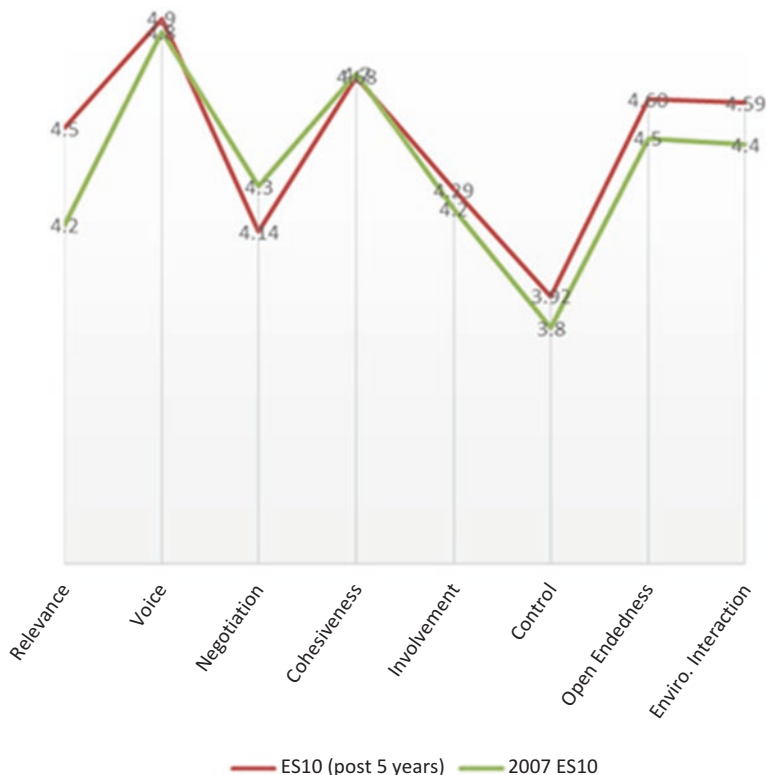


Fig. 5.1 Comparison of ES 102007 perceptions and five years later

group reuniting after several years. The questions for the focus group were designed to provide insight on respondents' long-held perception of ES 10's learning environment factors that they perceived to have affected them most as they relate to active citizenship components. The open-ended questionnaire contained sections related to active citizenship components and professional pathways.

Other questions included demographic information about the level of education completed, employment history, professional memberships or certifications, volunteerism, affiliation, long held beliefs about high school experiences and participatory practices. Qualitative and quantitative methods were used to increase the validity and reliability of the study by triangulating the qualitative results with the quantitative results (Creswell & Plano Clark, 2007). Data collected through the open-ended questionnaire and group interview were systematically analyzed through routine procedures to include traditional procedures using Microsoft excel and later using the qualitative software NVivo. The NVivo program helped organize the data beyond traditional approaches by sorting the coded data and making it easier to provide searches and cross referencing as well as frequency counting. This qualitative methodology was well suited to determine ES 10 graduates' perceptions

Table 5.8 Characteristic event: Volunteerism

Characteristic event	Notes
Volunteerism	<p>All graduates volunteered at the community level or beyond. 33% of graduates coached a sport 67% of graduates' volunteer in the community at various levels (environmental organizations, youth groups, outreach initiatives, church groups, homeless initiatives, poverty initiatives, women's shelters, medical initiatives, Social justice, youth engagement in democracy) 20% of graduates' volunteer in global initiatives (Red Cross, Africa Canada Accountability Coalition, OXFAM, Houses without borders, Global Health Initiatives, International Aid Worker)</p> <p>Note: Fifteen graduates reported getting involved with school organizations related to active citizenship in their grade 11, 12 years. Out of these fifteen, fourteen reported actively engaged in social justice, humanitarian, health, or environmental themed initiatives at the time of the study (6–7 years after completion of high school).</p>

toward lasting effects relating to active citizenship and linking these to learning environment features that students perceived as important. Table 5.9 demonstrates how aspects of the learning environment related to the PLACES inventory and how these aligned with outcomes of active citizenship as defined in the literature.

In summary of the ISSP survey results, the graduates of the ES 10 program demonstrated a high level of engagement in activities and initiatives that fit within the definition of active citizenship as proposed and conceptualized in this study. When compared to their Canadian counterparts, ES 10 graduates scored higher in most of the ISSP Citizenship 2004 survey (ISSP, 2012) categories. Based on a paired *t*-test, the differences in three of the categories were statistically significant. The three categories that were found to be significant were (1) Social and Political Action, (2) Good Citizen (measures community participation) and (3) Voice. Further the qualitative data from this study found that the ES 10 graduates indicated various forms of involvement in their communities, a result that was a strong indication that they were currently engaged in a varied level of active citizenship. All of the ES 10 graduates in the study volunteered in their community or beyond. Table 5.8 provides a summary of the various volunteerism reported by the ES 10 graduates.

4 Discussion

One of the sub questions in the study asked whether alumni believed that ES 10 had affected their civic engagements. Exploration of the participant responses was extended by probing to discover which particular activities, experiences or features of the ES 10 experience were seen as being important to the development of their civic engagement. Thus, this question provided a good opportunity to identify key learning environment features that the graduates described as having affected their civic engagement. Table 5.9 is intended to show connections between elements of the PLACES learning environment construct to active citizenship outcomes as

Table 5.9 Comparison of places constructs with active citizenship

Related PLACES scale and description	PLACES construct example from ES 10 participants	Active citizen related outcomes as demonstrated in the review of literature
<p>Relevance/integration: Extent to which lessons are relevant and integrated with environmental and community- based activities</p>	<p><i>The program showed us concrete examples of community commitment and activism. In university, I founded the non-profit organization: Africa Canada Accountability Coalition (Sarah)</i></p>	<p>An active citizen embraces social responsibilities and takes it upon themselves to play a civic role of being informed, maintaining and developing critical perspectives while becoming actively involved in social, political and/or environmental issues (Kincheloe, 2005).</p>
<p>Critical Voice: Extent to which students have a voice in the classroom procedures or protocols.</p>	<p><i>ES allowed me to voice my opinion ... coping with ambiguity and decision making in the classroom helped me to work with others in the future. Today I am confident in using my voice and self-advocacy which is important in my field of study (Marine Biology). (Lucas)</i></p>	<p>Empowerment and “giving people a voice” as well as taking responsibility and leadership. (European Commission Directorate General for Education and Culture (2007).</p>
<p>Student Negotiation: Extent to which students can negotiate activities in their class</p>	<p><i>We had a say in our learning which (then) led to cooperation and the acceptance of differences for the benefit of the group. I believe this has helped me with tolerance today when meeting people outside (my) usual crowd (Lily)</i></p>	<p>Important skills and attitudes related to active citizenship; Communication skills, debating skills, active listening skills, problem solving skills, coping with ambiguity, working with others and openness to change/ difference of opinion (Hoskins, 2006, p. 7).</p>
<p>Group Cohesiveness: Extent to which the students know, help and are supportive of one another</p>	<p><i>We were successful at creating a strong internal community.... This made a very strong impression on how important a support network is in life (Mike) ES encouraged a sense of caring for each other and the greater community. (Sharon)</i></p>	<p>Linking experience to opportunity; young people made connections between their opportunities and active citizenship experiences in various contexts (Ireland et al., 2006).</p>

(continued)

Table 5.9 (continued)

Related PLACES scale and description	PLACES construct example from ES 10 participants	Active citizen related outcomes as demonstrated in the review of literature
<p>Student Involvement: Extent to which students have attentive interest, participate in discussions, perform additional work and enjoy the class</p>	<p><i>To this day, I believe that ES10 was an innovative and engaging program that allowed students to not only learn through activities but also encouraged students to explore their natural curiosities in life and find something to care about. (Alex)</i></p>	<p>Linking experience to opportunity; young people made connections between their opportunities and active citizenship experiences in various contexts (Ireland et al., 2006).</p>
<p>Shared control: Extent to which teacher gives control to the students with regard to curriculum/activities</p>	<p><i>I remember appreciating the decision-making powers that our instructor granted us, and feel that the trust he placed within our group allowed us to achieve some things well beyond our years at the time... I believe we should have democratic control as to how we learn and work. This is instilled in ES. (David)</i></p>	<p>Having a voice; young people believed that they should have a voice on matters that affect them especially at school (Ireland et al., 2006).</p>
<p>Open Endedness: Extent to which the teacher gives freedom to students to think and plan own learning</p>	<p><i>Big one for me was the freedom of creativity, the flexible structure allowed the ability for one to expand on one's creative outlet. Coping with ambiguity was difficult but helped in critical thinking and decision making... Being pushed out of our comfort zone, helps in today's challenges. (Celeste)</i></p>	<p>Creativity, critical thinking skills, coping with ambiguity and informed decision making (Hoskins, 2006). Providing students opportunities to plan and implement actions that address real environmental problems in local communities is a powerful way of enhancing civic literacy (Orr, Strapp et al. in McClaren & Hammond, 2005).</p>
<p>Environmental interaction: Extent to which students are engaged in field or community-based experiences</p>	<p><i>ES helped me desire to better the world from an environmental perspective, through all the outdoor experiences and seeing what nature was all about. ES planted a seed to give to the greater community, to think outside yourself. (Emily)</i></p>	<p>Student involvement in place-based activities and communities of practice helps foster social and environmental action and responsibility (O'Connor & Sharp, 2013).</p>

described in the literature through illustrated examples how some alumni perceived the effects of particular program features and experiences on their current citizenship and community-related activities. For example, Sarah's comment (Table 5.9, Row 1) aligns with the PLACES construct of relevance and integration is connected to various activities that she recalled as occurring during the extended field experiences. Emily's comment (Table 5.9, Row 8) on the importance of being immersed in outdoor settings as a means to understand environmental issues as a key feature in her willingness to contribute aligns with the PLACES construct of environmental interaction and connects to the ES 10 goal of developing skill and knowledge in a range of field studies and outdoor pursuits. Both examples demonstrate how being immersed in community-based experiences can foster important beliefs and attitudes leading to active citizenship, which is consistent with the literature as illustrated in (Table 5.9, Column 3).

From the perspectives of Sarah and Emily, these two learning environment features were very important contributors to the development of their adult civic engagement. Further exploration into the responses from the graduates indicated the importance of how accepting and open they perceived the ES 10 learning environment to be. Sharon (Table 5.9, Row 4) believed ES 10 "*encouraged a sense of caring for each other and the greater community.*" She later spoke to this point during the consensus gathering part of the group interview, and her comments met with agreement from all other graduates. This group interview method included a consensus portion where common themes or outliers relating to the questions were identified by groups of graduates and then presented for all participants to determine if everyone was in agreement or had other points to add. Sharon's statement was as follows:

We were in grade 10 but felt we could have a big impact.... We learned to push ourselves further than ever before, everyone was pushing themselves, so it felt natural to do so.
(Sharon)

Sharon used the term "we" demonstrating that she felt comfortable describing this experience from a collective rather than individual perspective. Interestingly, many other responses from the group interview and questionnaires yielded similar responses referring to this collective experience using words like "us" and "we."

Another important piece from Sharon's earlier statement (Table 5.9, Row 4) is the importance of a "*sense of caring for each other and the greater community,*" which demonstrates the program fostered personal and social responsibility. Further, Sharon's comments above on how natural it was for students to push themselves in a collective way appear to recognize that although they were only in Grade 10 they were capable of much more than they might have expected from themselves.

It is important to note that a stated goal of the ES10 program was the development of "Friendships and positive peer relationships", and this connects to the PLACES construct of Group Cohesiveness: "Extent to which the students know, help and are supportive of one another." Being part of a strong sense of community where students trust and support each other is supported by the literature as a key feature to foster active citizenship as illustrated (Table 5.9, Row 4). What Sharon is

describing can be termed a community of practice. The concept of community of practice is attributed to the works of Lave and Wenger (Farnsworth et al., 2016). The key premise behind communities of practice is that they reflect fundamentally on the social nature of learning, which is illustrated when a group of people share a common concern or passion for something they do and go through a learning process together. When a community of practice develops, it also enables the social construction of knowledge. This learning takes place through shared experiences and co-participation in multiple learning practices such as those designed in a program such as ES 10. The following statement made by a graduate during the group interview phase of this research demonstrates participants' perception of the shared experience:

It was a crucial development point in our youth, we were allowed to experiment in a safe environment. Personal development through exploration grew to have strength in self which lead to sense of responsibility. There were demonstrated tangible benefits to include: communities based on values, personal growth, and a support network based on mutual trust developed skills leading to higher level of confidence and belief in oneself. Being responsive and taking responsibility was encouraged. We met people in the community which taught us skills and the importance of being involved. Experiencing small communities like on the Vancouver Island trip helped us realize that relationships were based on shared values rather than proximity. Working through real-life problems with community members gave us something to care about. (Peter)

It was noted that Peter's comments also met with consensus among the participants in the group interview session. What Sharon's and Peter's comments provide is a sense of what they believe to be the elements of ES 10 that may also have been important in fostering their community involvement following completion of the program. James uses the term "value" more than once in his comment. According to Rath et al. (1978), values are attitudes about the worth or importance of people, concepts or things. Values influence behaviour because one uses them to decide between alternatives. Values along with attitudes, behaviors and beliefs are foundational of who individuals are and how they do things (Rath et al., 1978).

Raths (as cited in Rath et al., 1978) focused on the process of valuing rather than values as being something static or fixed, which involved prizing one's beliefs, choosing one's beliefs and behaviours and acting on one's beliefs. The term value was used by many other students as well when describing their ES 10 experiences in relation to their interest and/or belief of making a difference in their communities, which aligns with Rath's valuing process. The influence of program experiences on value development is demonstrated by the following comment: "*The beach surveys (looking at change to our environment) and all the other outdoor experiences created a value and importance for the environment*" (Gerald). From the following graduate's perspective, shared values were prompted by "*the connection between the class and community helped realize your role as a citizen, there was a collective social responsibility here. The beach cleanup activity that we organized outside school time – was 100% initiated by us*" (Kerry). It is possible that shared values prompted by field experiences (attached to real-life problems) ignited a sense of agency in many students as illustrated by Kerry's comment.

A critical element here is that the sense of community that was established through classroom initiatives and to a larger extent through extended field experiences that allowed students to experience real-life phenomena issues and activities in local communities. In this heightened sense of community, students' perceptions of group cohesion were raised, as evident from their responses on the PLACES questionnaire and supporting qualitative data. Group cohesion is high when the "sense of caring" (Candice) can develop and when students are involved in experiential learning experiences centered around "real-life problems with community members" (Peter). Further, Peter saw high group cohesion as allowing students "to experiment in a safe environment," which was believed led to "personal development."

In addition, group cohesion translated to "being responsive and taking responsibility" because a "support network based on mutual trust" was built through experiences such as the one on Vancouver Island as referenced by Peter. The Vancouver Island experience included field experiences that saw the ES 10 students working collectively with community members and professional biologists to engage with a variety of real-life environmental issues. The trip was one week in duration wherein the class visited various communities and got involved in a wide range of activities. Examples of activities on the Vancouver Island trip included wetlands studies, foreshore and intertidal studies, forestry studies and land use studies. These investigations grew out of the concern of local community members. The following statement by Sue which met consensus during the group interview, which referred to these experiences on Vancouver Island, support Peter's claim: "This community involvement opened the idea of social responsibility ... we developed an appreciation of place and people developed through community interaction." The experiences gave ES 10 students something common to care about and may in turn have led to the community of practice effect seen in the students' descriptions.

ES 10 experiences appeared to have led to a heightened willingness for individual students to make contributions of sorts to their own communities. Emily's comment (Table 5.9, Row 8) supports this claim as she believed, "ES planted a seed to give to the greater community." It is important to note that the activities described on the Vancouver Island trip are consistent with the activities referred to by Sarah, Alex and Emily (Table 5.9, Rows 1, 5 and 8 respectively).

Further, collective groups of students from both the 2003 and 2004 cohorts reported involvement and collective contributions with volunteer organizations such as Stream Keepers and the Salmon Club while still in the ES 10 program and with volunteer organizations such as IMPACT (school group focusing on social justice issues), Juvenile Diabetes Research Foundation, The Salmon Club and Red Cross during their Grade 11 and 12 years. Many of these graduates attributed their experiences in ES 10 as stimulating their direct involvement in these programs, as evident by the following graduates comment:

There is no doubt in my mind that my grade 10 ES class allowed me to build a foundation of personal values that are based on a healthy natural environment and vibrant community. Following ES (while she was still in high school), I was asked to be the President of the leadership group, IMPACT. This volunteer group also allowed me to synthesize my passion for social justice. These two things encouraged me to find a degree to help influence in social justice. (Kerry)

Another common theme from the ES 10 alumni was the idea that the program contributed directly to their desire for and belief that they could make a difference by getting involved in community activities. A major finding of this study was that those students who got involved in volunteering through school opportunities provided while they were in their Grade 11 and 12 years were also more likely to continue volunteering in areas such as those relating to social justice, humanitarian, health or environmental themes after completion of high school. In fact, 14 of the 15 graduates who reported volunteering in school opportunities while in their Grade 11 and 12 years continued volunteering in their adult life in those areas mentioned. Further, 11 of the 15 graduates just mentioned expanded their involvement beyond the local community level to include involvement in global initiatives as well.

A major point to note is that while it appears the student's desires to get involved in active citizenship were ignited by the ES 10 program those who did continue to be involved in their Grade 11 and 12 years for the most part volunteered in school-supported initiatives such as Red Cross, IMPACT and the Salmon Club, and they did this collectively in small groups with fellow ES 10 students. In addition, since these graduates collectively participated with fellow ES 10 students in the mentioned initiatives, this indicates the importance of working with peers of similar interests.

Schools can play a role in the development of citizenship, and school environments can provide safe and supportive stepping stones or scaffolds into citizenship-related activities. These conditions can extend and complement the initiatives begun in programs such as ES 10. An important difference is that in ES 10, citizenship activities were developed as part of the core curriculum of the program, while the citizenship opportunities in Grades 11 and 12 were part of the EXTRA-curriculum. The "regular traditional" academic classes have learning environments that are not as supportive as ES 10 of this sort of active community involvement. If the development of citizenship is a core goal or mission of public schools, it is important to encourage practices and experiences in the regular curriculum that extend or are supportive of that mission rather than leaving it to chance or relegating it to the extra curriculum.

The educational model (Fig. 5.2) represents key learning environment features that can help foster the development of active citizenship. This model represents key learning environment features that can help foster the development of active citizenship indicators leading to long-term participatory action. Cohesive learning environments can be enhanced by team building and trust initiatives as well as integrated curriculum and flexible schedules which encourage prolonged engagement in collaborative learning activities. Learning environments high in group cohesion can be more successful when decisions are shared between the teacher and students around curriculum and schedule. Students that have an opportunity to exercise their voice regularly in open learning environments while participating collaboratively in various experiential learning opportunities that are community based can lead to self-discovery through active reflection while developing various skills, beliefs, attitudes, and values all related to being an active citizen. Those that continue their involvement in volunteering opportunities based on their new beliefs and desires may demonstrate a greater range of involvement in active citizenship.

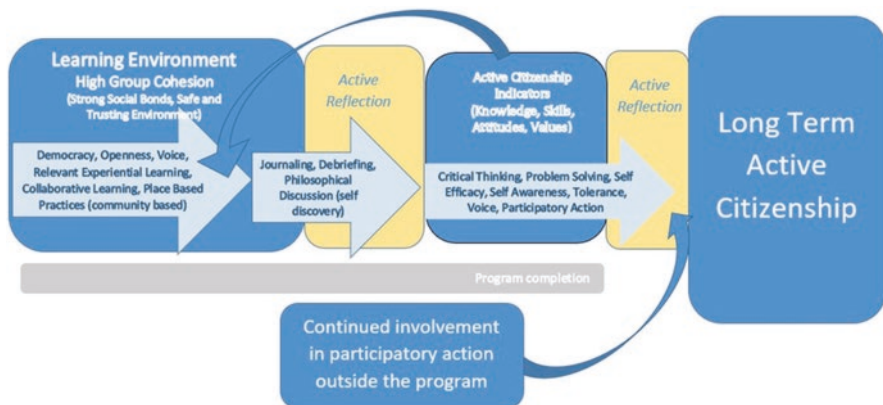


Fig. 5.2 Educational model for active citizenship (important learning environment features)

5 Limitations

This study was designed to investigate long-term effects of an ICP, Experiential Studies 10, on the development of active citizenship and to gain understanding of key learning environment features leading to this. The study is intended to help guide the development and implementation of educational programs with similar intents. With this in mind, several limitations must be acknowledged, and all claims and generalizations should be tempered by this knowledge. Member checking, peer debriefing and triangulation methods were utilized to minimize these concerns. Group interviews, although effective for gathering rich data, can also include the tendency for certain types of socially acceptable opinions to take form and permit certain individuals to dominate the process (Smithson, 2000). To address this limitation, Chartier's (2002) interview matrix method was used, which utilized smaller group interviews around the same questions and a consensus gathering portion. Finally, demonstrating the persistence of the PLACES survey by comparing the 2007 ES cohort's results with Koci's (2013) results helps increase the confidence in the participants' responses around the PLACES survey since this was perception based.

6 Importance of the Study

Research on learning environments, environmental learning and citizenship outcomes is still in its infancy. This study yields some interesting insight into the unique learning environments experienced by students in place-based education settings and has led to the increasing value of the PLACES instrument in the evaluation of learning environments in integrated programs. In the reported case study, students

noted a closer fit between their actual and preferred environments and often rated these settings more positively on all scales measured. This result also acknowledges the validity of the PLACES questionnaire over longer temporal timeframes, further strengthening its potential use as an evaluative tool for place-based and constructive learning environments. The PLACES questionnaire offers possibilities for studies in place-based environmental education settings, and offers new models for participatory action research by environmental educators. This opens up opportunities for future research to predict and describe other desirable learning outcomes that may prove to be associated with the learning environment facilitated in these programs. This was demonstrated with the ES 10 program where a very important learning feature of the program was how much say they had in everything, an attribute that they believed contributed to self-discovery and to caring about their learning experience. Democracy extended into the classroom can lead to self-determination where a student's voice is equal to that of the teacher's on many levels (Crittenden & Levine, 2016). Through place-based practices environmental programs like the one included in this study have demonstrated long term outcomes of active citizenship (Sturrock, 2017). This is just a small example of how a deeper understanding of learning environments in a place-based context can help environmental educators create more intentional experiences and more robust learning outcomes.

References

- Breunig, M., & Sharpe, E. (2009). Sustaining environment pedagogy in times of educational conservatism: A case study of integrated curriculum programs. *Environmental Education Research, 15*(3), 299–313. <https://doi.org/10.1080/13504620902807543>
- Chartier, R. (2002). *Tools for leadership and learning: Building a learning organization* (3rd ed.). National Managers' Community.
- Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. Sage.
- Crittenden, J., & Levine, P. (2016, Winter). Civic education. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from <https://plato.stanford.edu/archives/win2016/entries/civic-education/>
- Dorman, J. P., Fisher, D. L., & Waldrip, B. G. (2006). Learning environments, attitudes, efficacy and perceptions of assessment: A LISREL analysis. In D. L. Fisher & M. S. Khine (Eds.), *Contemporary approaches to research on learning environments: Worldviews* (pp. 1–28). World Scientific.
- Durr, K. (2004, December). *Education for democratic citizenship: The school a democratic learning community: The all European study on pupil's participation in school*. Council of Europe. Retrieved from http://www.coe.int/t/dg4/education/edc/Source/Pdf/Documents/2003_23_All-EuropStudyChildrenParticipation_En.PDF
- European Commission. (2007). *Study on active citizenship education*. Directorate-General for Education and Culture. Retrieved from http://ec.europa.eu/education/pdf/doc248_en.pdf
- Farnsworth, V., Kleanthous, I., & Wenger-Trayner, E. (2016). Communities of practice as a social theory of learning: A conversation with Etienne Wenger. *British Journal of Educational Studies, 64*(2), 139–160. <https://doi.org/10.1080/00071005.2015.1133799>
- Fisher, D. L., & Khine, M. (Eds.). (2006). *Contemporary approaches to research on learning environments: Worldviews*. World Scientific Publishing.

- Fraser, B. J. (1998). Science learning environments: Assessments, effects and determinants. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education* (pp. 527–564). Kluwer Academic Publishers.
- Fraser, B. J. (2007). Classroom learning environments. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 103–125). Routledge.
- Fraser, B. J. (2014). Classroom learning environments: Historical and contemporary perspectives. In N. Lederman & S. Abell (Eds.), *Handbook of research on science education, Volume II* (pp. 104–119). Routledge.
- Fraser, B. J., & Butts, W. L. (1982). Relationship between perceived levels of classroom individualization and science-related attitudes. *Journal of Research in Science Teaching*, 19(2), 143–154.
- George, D., & Mallery, P. (2003). *SPSS for windows step by step: A simple guide and reference* (4th ed.). Allyn & Bacon.
- Gruenewald. (2003). The best of both worlds: A critical pedagogy of place. *Educational Researcher*, 32(4), 3–12.
- Horwood, B. (1994). Integration and experience in the secondary curriculum. *McGill Journal of Education*, 29(1), 89–102. Retrieved from http://webapp1.dlib.indiana.edu/virtual_disk_library/index.cgi/4273355/FID1736/journals/enc2321/2321.htm
- Hoskins, B. (2006). *Draft framework on indicators for active citizenship*. CRELL. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.132.1723&rep=rep1&type=pdf>
- Hutchison, D. (2004). *A natural history of place in education*. Teachers College Press.
- International Social Survey Program Research Group. (2012). International Social Survey Programme: Citizenship – ISSP 2004. GESIS Data Archive, Cologne. ZA3950 Data file Version 1.3.0. <https://doi.org/10.4232/1.11372>.
- Ireland, E., Kerr, D., Lopes, J., & Nelson, J. (2006). *Active citizenship and young people: Opportunities, experiences and challenges in and beyond school citizenship education longitudinal study: Fourth annual report. Research report RR732* (Eric Document Reproduction Service No. ED502417). National Foundation for Educational Research. Retrieved from <http://files.eric.ed.gov/fulltext/ED502417.pdf>
- Johnson, D., & Johnson, R. (2003). *Cooperative learning: The social and intellectual outcomes of learning in groups* (Social Interdependence Theory) (pp. 136–176). RoutledgeFalmer.
- Kincheloe, J. (2005). *Critical pedagogy*. Peter Lang.
- Koci, P. (2013). *Factors influencing learning environments in an integrated experiential program*. Doctoral dissertation, Simon Fraser University, Burnaby, BC, Canada.
- McClaren, M., & Hammond, B. (2005). Integrating education and action in environmental education. In M. J. Mappin & E. A. Johnson (Eds.), *Environmental education and advocacy: Changing perspectives of environmental education* (pp. 267–291). Cambridge University Press.
- O'Connor, K., & Sharp, R. (2013). Planting the science seed: Engaging students in place – Based civic actions. *European Scientific Journal*, 4, 161–167. Retrieved from <http://ejournal.org/index.php/esj/article/view/2469/2342>
- Orr, D. (1992). *Ecological literacy*. State University of New York Press.
- Orr, D. (1994). *Earth in mind*. Island Press.
- Pickett, L., & Fraser, B. (2010). Creating and assessing positive classroom learning environments. *Childhood Education*, 86(5), 321–326. <https://doi.org/10.1080/00094056.2010.10521418>
- Raths, L. E., Harmin, M., & Simon, S. B. (1978). *Values and teaching* (2nd ed.). Charles E. Merrill.
- Revelle, W., & Zinbarg, R. (2009). Coefficients alpha, beta, omega, and the GLB: Comments on Sijtsma. *Psychometrika*, 74(1), 145–154. Retrieved from http://hbanaszak.mjr.uw.edu.pl/TempTxt/RevelleZinbarg_2008_ReliabilityCoefficients.pdf
- Smithson, J. (2000). Using and analyzing focus groups: Limitations and possibilities. *International Journal of Social Research Methodology*, 3(2), 103–119. <https://psycnet.apa.org/doi/10.1080/136455700405172>
- Sobel, D. (1993). *Children's special places*. Zephyr Press.
- Sobel, D. (1996). *Beyond ecophobia: Reclaiming the heart in nature education*. Orion Society.

- Sturrock, G. (2017). *Long-term perceived outcomes of an integrated curriculum program as it relates to active citizenship*. Doctoral dissertation, Simon Fraser University, Burnaby, BC, Canada.
- Tri-Cities Chamber of Commerce. (2014). Economic profile – Coquitlam/Port Coquitlam/Port Moody. http://www.coquitlam.ca/docs/default-source/why-coquitlam-documents/Tri-Cities_Chamber_of_Commerce_Economic_Profile_2014.pdf?sfvrsn=0
- Thomashow, M. (1996). *Ecological identity*. MIT Press.
- Woodhouse, J., & Knapp, C. (2000). *Place-based curriculum and instruction* [ERIC Document Reproduction Service No. EDO-RC-00-6].
- Zandvliet, D. B. (2007, November). *Learning environments for environmental education*. Paper presented at the Australian Association for Research in Education (AARE), Fremantle, Australia. Retrieved from <http://www.aare.edu.au/publications-database.php/5546/learning-environments-for-environmental-education>
- Zandvliet, D. B. (2012). Development and validation of the place-based learning and constructivist environment survey (PLACES). *Learning Environments Research*, 15(2), 125–140. <https://doi.org/10.1007/s10984-012-9110-x>
- Zandvliet, D. B. (2014). PLACES and SPACES: Case studies in the evaluation of post-secondary, place-based learning environments. *Studies in Educational Evaluation*, 41, 18–28. <https://doi.org/10.1016/j.stueduc.2013.09.011>
- Zandvliet, D. B., & Fraser, B. J. (Eds.). (2018, October). *Thirty years of learning environments: Looking back and looking forward*. Brill/Sense.

Gordon Sturrock is a coordinator and faculty member of Douglas College in the Sport Science Department, faculty of Science and Technology. He teaches courses that center around pedagogy, physical literacy, and alternative environments. He has vast teaching experience especially within experiential education programs that spans the elementary, secondary, and post-secondary levels of education. His career interests lie in pedagogical practices within a variety of contexts relating to learning environments and long-term effects of active citizenship.

David Zandvliet is a Professor and UNESCO Chair at Simon Fraser University in Vancouver, Canada and the founding Director for the Institute for Environmental Learning. An experienced researcher, he has published articles in international journals and presented conference papers on six continents and in over 15 countries. His career interests lie in the areas of science and environmental education and learning environments. He has considerable experience in the provision of teacher development and has conducted studies in school-based locations in the US, Australia, Canada, Indonesia, Malaysia, Sri Lanka and Taiwan.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 6

Fostering Effective Teaching at Schools Through Measurements of Student Perceptions: Processes, Risks and Chances



Hannah J. E. Bijlsma  and Sebastian Röhl 

Abstract Student perceptions of teaching quality have become increasingly important for measuring teaching effectiveness and can be used for the subsequent improvement of teachers' teaching. However, measuring teaching quality through student perceptions reliably and validly and the subsequent improvement is not guaranteed. On the one hand, students' teaching quality data are influenced by many characteristics of the students, classes and measurement instruments, and on the other hand, teachers' use of the feedback data is influenced by factors such as personality, context and data characteristics. This chapter, therefore, provides important insights into measuring teacher effectiveness through student perceptions, risks and opportunities of using these teaching quality perceptions and the effective use of student feedback data for the development of teaching and teachers.

Keywords Student perceptions · Teaching quality · Feedback · Teacher development

1 Introduction

Within schools, teaching quality is one of the most important factors in student achievement (Nye et al., 2004; Rivkin et al., 2005). Thus, in order to address the decline in student achievement all over the world (OECD, 2014), increased emphasis has been placed on examining teaching quality and improving teacher effectiveness (Timperley et al., 2007). Teaching quality can be determined in several

H. J. E. Bijlsma (✉)
Research Department, Dutch Inspectorate of Education,
Utrecht, The Netherlands
e-mail: h.j.e.bijlsma@owinsp.nl

S. Röhl
Institute for Educational Sciences, Eberhard Karls University Tübingen,
Tübingen, Germany
e-mail: sebastian.roehl@uni-tuebingen.de

ways; for example, through lesson observations by external observers to analyze student achievement growth, or by teacher self-evaluation. All of these approaches have their advantages and disadvantages.

In addition to the above-mentioned methods, student perceptions of teaching quality have become increasingly important for measuring teacher effectiveness (Bell & Aldridge, 2014; Ferguson, 2012; Goe et al., 2008). Students' ratings for a lesson can be used for conducting research on, for example, the effectiveness of classroom interventions, and, to a limited extent (see Part III), for accountability purposes at schools. Moreover, with the student ratings, teachers can identify where improvement of their teaching is still possible and they can make their teaching more effective for student learning (Gärtner, 2014; Peterson et al., 2000). Student perceptions are thus considered very helpful for developing instructional quality. For example, in the early years of teacher effectiveness research, Gage (1960) studied sixth grade teachers receiving information as to how their students described their actual and their ideal teacher. More recently, Bell and Aldridge (2014) investigated the use of student perception data for teacher reflection and classroom improvement, and Mandouit (2018) used action research to investigate the impact of student feedback on teacher practices. A recent meta-analysis of student feedback intervention studies was able to show that, on average, the use of student feedback on teaching can indeed generate a significant, albeit small, positive effect on teaching quality as viewed from the student's perspective (Röhl, 2021). Notably, the systematic literature search for this meta-analysis revealed that, with the exception of one study from Turkey, only intervention studies from Western countries were found, even though student perceptions are assumed to be as effective for measurements of teaching quality and learning environments in Eastern countries and cultures as well (e.g., Khalil & Aldridge, 2019; Maulana et al., 2012).

Some issues have been raised concerning the reliability and validity of students' perceptions for assessing teaching quality. Various statistical techniques can be used to correct for these problems, namely, Classical Test Theory, Item Response Theory or Generalizability Theory. These techniques function as being exemplars for the connection between psychometric theories and the different perspectives on the validity of student perceptions (Bijlsma et al., 2021).

However, the arguments for and against the use of student ratings as a basis for improving teaching have been going on for some time now. And even if student ratings were guaranteed to be accurate measures of teaching quality, the ratings cannot in themselves support improvement of individual teaching performance (Loeb, 2013). For improvement to occur, it is also necessary for teachers to meaningfully reflect on the feedback they receive and use it to develop and implement improvement-oriented actions.

Therefore, in this chapter, we first present a process model of the use of student feedback in schools that visualizes its productive use for the improvement of teaching quality. This model illustrates that, on the one hand, the teaching quality data are influenced by several characteristics of the students, classes, and measurement instruments, and, on the other hand, teachers' use of the feedback data is influenced by factors such as personality, context and data characteristics. The advantage of

this model lies in its cyclic way of looking at student feedback utilization by teachers, instead of a linear approach, used, for example, by Gärtner (2014), and which further does not consider factors influencing students' perceptions and feedback. Following this, we present an overview of the empirical literature on peculiarities of student perception data, especially concerning validity, reliability and potential factors influencing student ratings, and discuss how these measurement characteristics should be considered by teachers when using student ratings of teaching quality for the improvement of their teaching. This is followed by an overview of factors influencing the utilization of student feedback for the improvement of teaching and teachers. Lastly, we consider the conditions under which teachers' process of collecting, interpreting and accepting the data, and subsequent teaching improvement can be accomplished. Opportunities for further research are presented.

In this chapter, thus, we give an overview of the literature, focussing on what we know about student feedback on teaching and what teachers should keep in mind when they perceive and utilize the feedback for their professional development and improvement of teaching. With this overview, we aim to provide important insights into measuring teacher effectiveness through student perceptions, risks and opportunities of these teaching quality perceptions, and the effective use of student feedback data for the development of teaching and teachers.

2 Process Model of Student Feedback on Teaching

The process of using students' teaching quality ratings to improve instructional quality has many necessary stages and is influenced by many individual and contextual factors, starting with the specifics of obtaining information about teaching quality using student perception questionnaires. To make sure that the information available in the teaching quality data actually leads to professional development of teaching, the teachers must transform the information into improvement-oriented actions. Such actions include giving special attention to possible areas of improvement during lesson preparation or teaching, attending targeted training courses, asking colleagues for advice, or looking for ways to improve the teaching situation together with the students (for an overview, see Röhl, 2021; Bijlsma et al., 2019b). Unfortunately, receiving feedback does not automatically lead to improvement processes. Röhl et al. (2021) summarized findings from organizational psychology on productive feedback use (Ilgen et al., 1979; Kahmann & Mulder, 2011; Kluger & DeNisi, 1996; Smither et al., 2005) in a model to visualize teachers' feedback use processes (Fig. 6.1).

Once the feedback information is available, the teacher has to perceive, understand, and interpret the data. Teachers need a form of data literacy (Kippers et al., 2018; Mandinach & Gummer, 2013) to interpret the information in feedback reports correctly. Additionally, reactions to received feedback have not only cognitive, but also affective components (Kahmann & Mulder, 2011; Taylor et al., 1984). Therefore, during this interpretation process, positive emotions such as satisfaction

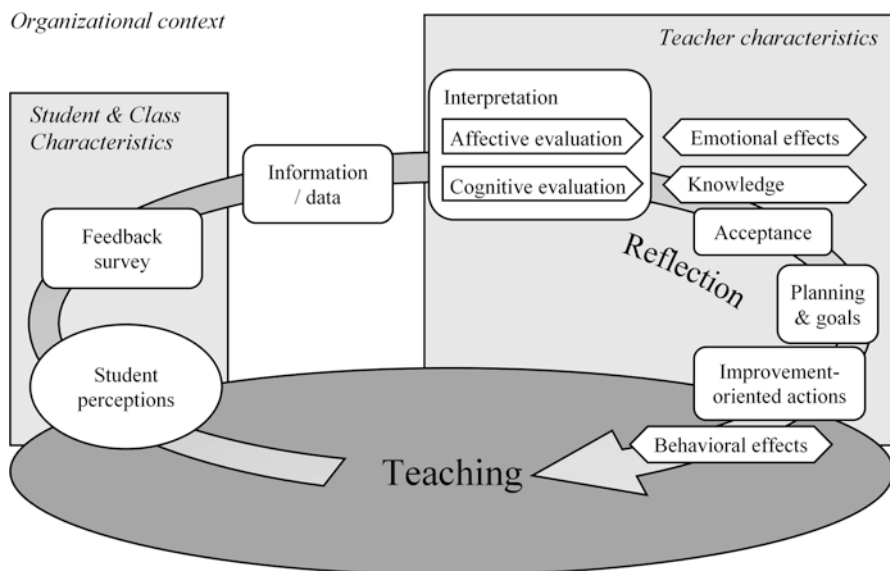


Fig. 6.1 Process model of student feedback on teaching. (Source: Röhl et al., 2021, p. 4)

and joy, or negative ones such as dissatisfaction or defensiveness can occur as emotional effects. On the cognitive level, knowledge effects can occur when feedback provides the teacher with new information about the students' view of their teaching or the feedback reinforces their existing knowledge.

The new knowledge is linked to the teacher's own perceptions and standards for teaching. Any discrepancies must be considered (i.e., the feedback that contradicts one's own perceptions) in order for the teacher to consider changes in their teaching. This could lead to the teacher's planning and goal-setting for the elimination of a discrepancy in a possible area of improvement (Smither et al., 2005), which could finally result in improvement-oriented actions as behavioral effects of the feedback. This process on the part of the teacher represents, in a sense, the bottleneck for realizing the potential of student feedback for teaching improvement. This process is influenced by factors concerning the students and classes, the teacher, and the organizational context, the importance of which for the practice of student feedback use we discuss below.

3 Factors Associated with Student Perception Measurements

Perceptions of the quality of the same teaching practices differ between students. These differences are not undesirable per se, because ratings do reflect a student's personal perspectives on teaching quality, and students do differ (Kenny, 2004).

Insight into the extent to which differences in student ratings are related to factors on the student, teacher and class levels is important for evaluating the ratings students give and avoiding any incorrect conclusions. For example, the average teaching quality score can be lower in a class with many low-performing students without the teaching quality actually being lower. Female teachers might receive significantly lower ratings from male students although they are doing as good a job as male teachers do. In the following section, we discuss factors associated with student perceptions of teaching quality on four levels: characteristics of students, teachers, classes and measurements.

3.1 Student Characteristics

Some research has reported that teachers at both the primary and secondary school levels were viewed as more dominant, more positive and more cooperative by girls than by boys (Den Brok et al., 2006; Fisher et al., 2006; Levy et al., 2003; Rickards, 1998; Veldman & Peck, 1969). However, it is not clear to what extent the gender effect is confounded with the effects of other variables, as gender seems to interact with a number of other variables, such as students' subject preferences (Baker & Leary, 1995; Jones & Kirk, 1990), ethnicity or culturally-related gender role definitions (Levy et al., 2003; Timm, 1999; Worthington, 2002) and level of academic performance (Brophy & Good, 1986; Goh & Fraser, 1995; Levy et al., 2003). Student age was found to be related to student perceptions of their teacher, as older students tend to perceive their teachers as more strict and noted more teacher dominance than their younger peers in some studies (Levy et al., 1997; Levy et al., 2003). Moreover, students with higher general interest in the subject are more likely to give a higher rating of teaching quality than students with lower interest (Cashin, 1988; Fisher et al., 2006). Students' achievement was also found to be related to their perceptions of their teacher: Students with high prior achievement tend to perceive the quality of their teacher's teaching more positively than students with low prior achievement (Atlay et al., 2019; Bijlsma et al., 2022; Gärtner & Brunner, 2018; Marsh, 2007). Additionally, the level of parental education and wealth of the students should be considered, as a study by Atlay et al. (2019) pointed towards a negative association of these characteristics with student perceptions of their teachers' behavior.

3.2 *Teacher Characteristics*

Mixed results have been found for teacher gender influencing student ratings of teaching quality. Veldman and Peck (1969) found a significant but weak effect of teacher gender, showing that female secondary school teachers tend to receive higher ratings than their male colleagues, but this effect was only found for being ‘friendly and cheerful’ and not for other aspects of teaching quality. Bijlsma et al. (2022) did not find any significant effects of gender on student ratings. They studied effects of teacher popularity on student perceptions of teaching quality and found that the more popular the teacher is according to their students, the higher students’ ratings of their teaching qualities. This relationship was also addressed by Gärtner (2014), Gärtner and Brunner (2018), Clausen (2002), Fauth et al. (2014), Goe et al. (2008) and Donahue (1994). In addition, teachers with more teaching experience receive higher teaching quality ratings from their students than teachers with little teaching experience (Bijlsma et al., 2022; Brekelmans et al., 2002; Day et al., 2008; Kini & Podolsky, 2016; Leigh, 2010; Rowley, 2003). Other variables mentioned in the literature that might influence student ratings of their teacher are teachers’ cultural and ethnic background, whereby teachers from another ethnic background than the student receives lower teaching quality ratings (den Brok et al., 2002; den Brok et al., 2003), teachers’ personality, whereby more stressed teachers are rated as less socially oriented (Klusmann et al., 2006), and teachers’ teaching ability or capacity, whereby lower ability or capacity results in lower teaching quality ratings (Veldman & Peck, 1969).

3.3 *Class Characteristics*

Compared to the student and teacher factors, less is known about class-level factors influencing students’ perceptions of teaching quality. Class size might be related to differences in student ratings, as teachers might have more difficulty with classroom management in large classes, which is reflected in the students’ teaching quality ratings. In a study by Levy et al. (2003), however, it appeared that class size was negatively related to student perceptions of teacher proximity and unrelated to their perceptions of teacher influence. According to Bijlsma et al. (2022), class size also did not matter for the students’ perception of teaching quality. However, according to Göllner et al. (2020), classes with higher proportions of boys and lower mean achievement levels had lower teacher scores for classroom management. Fisher et al. (2006) found that students in highly motivated classes had more favorable perceptions of their teachers. Moreover, they concluded that class composition variables such as percentage of students with a migration background seemed important for differences in student ratings (on average, those classes rated their teachers lower). Bijlsma et al. (2022) however, did not find an impact of the ethnic make-up of the class on students’ perceptions of teaching quality. Other class-level variables

that are related to student perceptions of teaching quality are the subject being taught by the teacher (Gärtner & Brunner, 2018; Veldman & Peck, 1969) and the class' average level of academic achievement (Bijlsma et al., 2022; Veldman & Peck, 1969).

3.4 Measurement Characteristics

Although a student perception questionnaire can be seen as text material in normal language (i.e., textual information presented in the form of separate items; Tourangeau et al., 2000), existing student perception questionnaires differ fundamentally in their linguistic complexity, which shapes student responses (Göllner et al., 2021; Krosnick & Presser, 2010; Tourangeau et al., 2000). It can therefore be argued that differences in student ratings of their teaching quality arise because students encounter difficulties in comprehending the questionnaire items. For example, items that include many linguistic features, including surface aspects (e.g., the length of words and sentences) and characteristics that require more linguistic analysis (e.g., the number of complex noun phrases) can be difficult to understand. Moreover, an item's referent (the subject to which an item refers) and addressee are two salient characteristics that might affect the information obtained from student ratings of teaching quality. Measurement characteristics also refer to the frequency of measurements (time between the assessments; Gärtner & Brunner, 2018) and to the anonymity of the ratings (Gärtner, 2014).

4 Interpreting and Analyzing Student Feedback Data

Insight into the factors related to differences in student perceptions of teaching quality as presented in Sect. 3 can strengthen the general awareness among teachers of the required nuanced and careful interpretation of student feedback (Bijlsma et al., 2022; Den Brok et al., 2006). For example, if a teacher receives high teaching quality ratings from their students, it is good to be aware that this could have to do with, for example, being a good teacher, popularity (for some reason), or the fact that there are many high-performing or highly motivated students in the class in question. In lower grades teachers' interpretation of very positive ratings regarding their teaching quality should be more cautious than in the higher grades, as teachers' proximity to younger students might be greater than their proximity to older students, which might cause a strong effect on teaching quality ratings. Of course, not all of the factors presented above always represent a bias in reported teaching quality. For example, it is to be expected that teachers with a higher level of experience will also have higher reported teaching quality, and that teachers with a high level of stress will find it more difficult to deliver lessons of a high quality.

In addition to gaining knowledge of the factors influencing student perceptions for the most valid interpretation of the feedback received, it is advisable for teachers to disclose the feedback received to the class. By doing so, the teacher can ask directly about specific conspicuous aspects and how these results are to be interpreted from the class's point of view. Although this may remove the veil of anonymity for student respondents, the information in the feedback can be exploited, for example, by identifying and clarifying misunderstandings of item formulations and other rating biases.

Scientific findings have indicated that not only the mean values, but also the consensus of students' ratings on teaching quality within classes is predictive for learning achievement (Schweig, 2016). Thus, if students' answers to an item differ strongly within a class, this can be seen as an important indication of possibilities for improving one's own teaching in this respect.

As called for in many places (AERA, APA, & NCME, 2014; Bell, 2019; Hill et al., 2011), the validity of student perception measures should always be considered in light of the purpose of data collection. The following situations can be distinguished: (a) teachers voluntarily searching for feedback on their own initiative, (b) student feedback delivered to teachers as established practice or given by the organization, but without official accountability purpose, and (c) student feedback with accountability purposes (Röhl & Gärtner, 2021). The interpretation and analysis of formative student feedback to teachers with the purpose of professional development must be clearly distinguished from any form of summative evaluation, assessment, or rating that is used for administrative decisions.

5 Relevant Conditions for Teachers' Utilization of Student Feedback

Careful interpretation of the student feedback data is included in the Process Model of Student Feedback on Teaching (presented in Sect. 2 of the chapter) by teachers' reflection and action phases and subsequent improvement of teaching quality. In order words, teachers may utilize the feedback data to work on improving their instruction.

Many findings and theories from feedback research point to the relevance of both individual teacher characteristics and organizational characteristics for teachers' use of student feedback for improving teaching quality. In this section, we will outline relevant factors influencing teachers' use of student feedback from both an organizational psychology perspective (Ilgen et al., 1979; Smither et al., 2005) and a data-based decision-making perspective (Brunner & Light, 2008; Schildkamp & Lai, 2013; Schildkamp et al., 2013).

5.1 Characteristics of Feedback Recipients (Teachers)

Empirical findings show that teachers' age and professional experience affect teachers' use of student feedback. In general, older teachers seek less collegial feedback (Kunst et al., 2018; Runhaar et al., 2010) and use feedback less often compared to younger teachers (Ditton & Arnold, 2004). Teachers with longer professional experience are more skeptical of the usefulness of feedback (Dretzke et al., 2015). Some findings on gender effects regarding feedback show that female teachers more often seek collegial feedback (Runhaar et al., 2010) and tend to improve their teaching more after receiving and utilizing student feedback (Buurman et al., 2018). Teachers with higher self-efficacy seek more feedback and are more willing to reflect upon it (Ditton & Arnold, 2004; Runhaar et al., 2010). Moreover, teachers' motivation to use the feedback data for improving teaching quality is a relevant factor (Bijlsma et al., 2019a), as well as teachers' data literacy (their ability to understand numerical or other data and translate them into actions; Mandinach & Gummer, 2016; Schildkamp et al., 2017). Other individual characteristics of teachers that might foster the processing and use of student feedback are high mastery goal orientation (Elliott & Dweck, 1988), lower level of perceived stress (Ditton & Arnold, 2004; Elstad et al., 2015), and more positive attitude towards students' trustworthiness or competence as feedback providers (Balch, 2012; Ditton & Arnold, 2004; Elstad et al., 2017; Ilgen et al., 1979).

5.2 Characteristics of the Organization (School)

A feedback culture is generally defined by different organizational characteristics, such as support for giving and interpreting feedback, a non-threatening atmosphere, shared valuing of feedback for improvement, team psychological safety, and support in understanding feedback, setting goals, and implementing them in practice. In general, a well-established feedback culture has proved to be effective for the use of feedback in organizations (London & Smither, 2002). In the context of student feedback, in particular, those intervention studies that provided supportive measures for reflection and teaching development showed significantly higher positive effects (Röhl, 2021). In all of this, leadership plays an important role in feedback usage processes (Röhl & Gärtner, 2021). In an educational setting, it is important that school leaders have a clear vision of the schools' future, inspire teachers in their work, give the work a greater sense of meaning, and stimulate the questioning of old assumptions (transformational leadership; Bass, 1985; Runhaar et al., 2010). Active encouragement by school leaders to seek student feedback is also supportive, as extrinsically motivated feedback use is as beneficial to reported improvements in teaching as is intrinsically motivated feedback use (Gärtner, 2014; Röhl & Gärtner, 2021). However, it is important to ensure that the use of feedback is communicated as an opportunity for development and not as control or accountability, as the latter

can lead to resistance to its use (Elstad et al., 2017). School leaders should also give teachers the feeling of autonomy to make decisions about their instruction in data-use processes in schools (Prenger & Schildkamp, 2018).

5.3 *Characteristics of Feedback Information (Data)*

With regard to the characteristics of the feedback message, the comprehensibility, valence, specificity and timing of the feedback data are relevant in the processing and use of feedback (Röhl & Gärtner, 2021). The feedback data need to be presented in such a way that teachers understand the results, for example, mean scores in graphs or scale plots, or means for every item. The more positive the feedback, the more precise reception, easier remembering of contents, and better acceptance of the feedback by teachers (Ilgen et al., 1979; Lyden et al., 2002). The literature shows different findings on the specifics of the feedback, ranging from ‘highly specific feedback’ to ‘low specificity or summarized feedback’. High-specificity feedback seems to be more effective for beginners and for short-term learning, whereas low-specificity feedback tends to have a stronger impact on long-term learning performance (Röhl & Gärtner, 2021).

The timing of the feedback refers to the time between the actual act or task and the provision of the feedback. If the feedback is provided to the teacher right after a lesson, the link between the actual actions of the teacher in the classroom and the student feedback is clearer than in the case of feedback on teacher behavior in general (across many lessons; Hattie & Timperley, 2007; Shute, 2008). When feedback is given immediately, it is found to be more effective than when it is postponed (Timmers & Veldkamp, 2011). Teachers might therefore be able to work better on improving their teaching quality when feedback is given immediately (Bijlsma et al., 2019b). Furthermore, a survey instrument that is scientifically and psychometrically validated and reliable should be carefully selected for reliable and valuable use of student feedback data (Bijlsma, 2021).

6 **Conclusions and Future Directions**

Student feedback can be a valuable tool to improve teaching. However, teachers’ use of feedback data to assist in their professional development does not happen automatically. On the basis of the Process Model of Student Feedback on Teaching (see above, Röhl et al., 2021), we pointed out that on the one hand, student teaching quality perceptions are influenced by several characteristics of the students, classes and measurement instruments, and on the other hand, teachers’ use of the feedback data is influenced by factors such as individual characteristics of the teacher, and context and data characteristics. Insight into these factors can strengthen the general awareness among practitioners of the conditions under which teachers’ process of

collecting, interpreting and valuing the results, and the subsequent teaching improvement, can be accomplished successfully.

For future research, an interesting question is how the prerequisites for teacher development based on student feedback can be fulfilled to match with what is possible within the context of schools. From the research on deliberate practice by professionals and experts by Ericsson (2006), we know that improving as a teacher requires a coach who guides the teacher through the improvement process and who knows what ideal teaching behavior looks like, how this behavior can be trained effectively, and what practices are effective if problems occur during the improvement process. From the research on Professional Learning Communities (e.g., Brown & Poortman, 2018), we know that teacher collaboration in improvement processes is a promising way to improve teachers' teaching, in which the underlying goal is to improve teaching and teacher learning within the school (Blankenship & Ruona, 2007; Prenger et al., 2017). We recommend investigating the role of a coach and the collaborative learning process among teachers when improving teaching quality based on student feedback.

Moreover, it would be profitable to investigate the use of student feedback data for improving teaching quality in non-Western cultures. Although student perceptions have mainly been used in Europe, Australia and the USA thus far, we assume that they might also be useful in non-Western school cultures. There are studies on student perceptions of teaching quality in schools and also on its use in higher education, for example in Asian countries (e.g., Maulana et al., 2012). However, to the best of our knowledge, there is a lack of studies dealing with how student perceptions of teaching quality can be used as feedback to teachers for the purpose of improving teaching in primary and secondary schools. Adapting findings from Western cultures to the cultural conditions in non-Western cultures might be necessary here.

Another direction for future research might be to combine different teaching quality measures (e.g., classroom observations, student perceptions and teacher perceptions) to obtain a rich picture of teaching quality. Some aspects of teaching quality, for example, are probably best assessed by students, such as whether students feel that the teacher has high expectations of them, and whether students experience the classroom climate as safe. To understand other teacher quality aspects, other perspectives might be more relevant. For example, does an external observer, based on his or her professional standards, think that the explanation of subject matter by the teacher is correct? Moreover, as far as teachers' perspectives on their lessons are concerned, it would be interesting to know how they perceive their own teaching quality and compare this with the student perceptions, as this may influence their opinion about the need for improvement of their lessons.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. American Educational Research Association.
- Atlay, C., Tieben, N., Fauth, B., & Hillmert, S. (2019). The role of socioeconomic background and prior achievement for students' perception of teacher support. *British Journal of Sociology of Education*, 40(7), 970–991. <https://doi.org/10.1080/01425692.2019.1642737>
- Baker, D., & Leary, R. (1995). Letting girls speak out about science. *Journal of Research in Science Teaching*, 32, 3–27. <https://doi.org/10.1002/tea.3660320104>
- Balch, R. T. (2012). *The validation of a student survey on teacher practice*. Doctoral dissertation, Vanderbilt University. Vanderbilt University Institutional Repository Electronic Theses and Dissertations. <https://etd.library.vanderbilt.edu/etd-07182012-153440>
- Bass, B. M. (1985). *Leadership and performance beyond expectations*. Free Press.
- Bell, C. A. (2019). *What we see depends on how we look* QUINT conference, Oslo.
- Bell, C. A., & Aldridge, J. M. (2014). Investigating the use of student perception data for teacher reflection and classroom improvement. *Learning Environment Research*, 17, 371–388. <https://doi.org/10.1007/s10984-014-9164-z>
- Bijlsma, H. J. E. (2021). The quality of student perception questionnaires: A systematic review. In W. Rollett, H. J. E. Bijlsma, & S. Röhl (Eds.), *Student feedback on teaching in schools. Using student perceptions for the development of teaching and teachers* (pp. 47–72). Springer. https://doi.org/10.1007/978-3-030-75150-0_3
- Bijlsma, H. J. E., Keuning, T., & Visscher, A. J. (2019a, Aug 10–14). *A users' perspective on the use of a digital tool for student feedback about teaching quality*.
- Bijlsma, H. J. E., Visscher, A. J., Dobbelaer, M. J., & Veldkamp, B. P. (2019b). Does smartphone-assisted student feedback affect teaching quality? *Technology, Pedagogy & Education*, 28(2), 217–236. <https://doi.org/10.1080/1475939X.2019.1572534>
- Bijlsma, H. J. E., van der Lans, R. M., Mainhard, M. T., & den Brok, P. (2021). A reflection on student perceptions of teaching quality from three psychometric perspectives: CCT, IRT and GT. In W. Rollett, H. J. E. Bijlsma, & S. Röhl (Eds.), *Student feedback on teaching in schools. Using student perceptions for the development of teaching and teachers* (pp. 15–30). Springer.
- Bijlsma, H. J. E., Glas, C. A. W., & Visscher, A. J. (2022). Factors related to differences in digitally measured student perceptions of teaching quality. [Paper accepted]. *School Effectiveness and School Improvement*, 33, 360–380.
- Blankenship, S. S., & Ruona, W. E. (2007). Professional learning communities and communities of practice: A comparison of models, a literature review. *Online submission*.
- Brekelmans, M., Wubbels, T., & den Brok, P. (2002). Teacher experience and the teacher-student relationship in the classroom environment. In S. C. Goh & M. S. Khine (Eds.), *Studies in educational learning environments: An international perspective* (pp. 73–100). New world scientific. https://doi.org/10.1142/9789812777133_0004
- Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. Witrock (Ed.), *The third handbook of research on teaching* (pp. 328–375). Macmillan.
- Brown, C., & Poortman, C. L. (2018). *Networks for learning: Effective collaboration for teacher, school and system improvement*. Routledge Taylor & Francis Group. <https://doi.org/10.4324/9781315276649>
- Brunner, C., & Light, D. (2008). From knowledge management to data-driven instructional decisionmaking in schools: The missing link. In A. Breiter, A. Lange, & E. Stauke (Eds.), *School information systems and data-based decision-making* (pp. 37–48). Peter Lang.
- Buurman, M., Delfgaauw, J. J., Dur, R. A. J., & Zoutenbier, R. (2018). *The effects of student feedback to teachers: Evidence from a field experiment*. Tinbergen Institute.
- Cashin, W. E. (1988). *Student ratings of teaching: A summary of the research* (IDEA Paper No. 20) (p. 29). Center for Faculty Evaluation and Development, Kansas State University.

- Clausen, M. (2002). *Qualität von Unterricht: Eine Frage der Perspektive?* [Quality of instruction as a question of perspective?]. Waxmann.
- Day, C., Sammons, P., & Kington, A. (2008). *Effective classroom practice (ECP): A mixed-method study of influences and outcomes*. School of Education, University of Nottingham.
- den Brok, P., Levy, J., Rodriguez, R., & Wubbels, T. (2002). Perceptions of Asian- American and Hispanic-American teachers and their students on interpersonal communication style. *Teaching and Teacher Education*, 18, 447–467. [https://doi.org/10.1016/S0742-051X\(02\)00009-4](https://doi.org/10.1016/S0742-051X(02)00009-4)
- den Brok, P., Levy, J., Wubbels, T., & Rodriguez, M. (2003). Cultural influences on students' perceptions of videotaped lessons. *International Journal of Intercultural Relations*, 27(3), 355–374. [https://doi.org/10.1016/S0147-1767\(03\)00016-6](https://doi.org/10.1016/S0147-1767(03)00016-6)
- Den Brok, P., Fisher, D. L., & Rickards, T. (2006). Factors affecting Australian students' perceptions of their science teachers' interpersonal behaviour. In D. L. Fisher & M. S. Khine (Eds.), *Contemporary approaches to research on learning environments: World views* (pp. 51–74). World Scientific.
- Ditton, H., & Arnold, B. (2004). Wirksamkeit von Schülerfeedback zum Fachunterricht [Effectiveness of student feedback on subject teaching]. In J. Doll & M. Prenzel (Eds.), *Bildungsqualität von Schule. Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsentwicklung* (pp. 152–172). Waxmann.
- Donahue, J. M. (1994). Student perceptions of their teachers, their school, and themselves as learners. *Retrospective theses and dissertations*.(10694), Retrieved from <https://lib.dr.iastate.edu/rtd>. <https://doi.org/10.31274/rtd-180813-9962>
- Dretzke, B. J., Sheldon, T. D., & Lim, A. (2015). What do K-12 teachers think about including student surveys in their performance ratings? *Mid-Western Educational Researcher*, 27(3), 185–206.
- Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, 54, 5–12. <https://doi.org/10.1037/0022-3514.54.1.5>
- Elstad, E., Lejonberg, E., & Christophersen, K.-A. (2015). Teaching evaluation as a contested practice: Teacher resistance to teaching evaluation schemes in Norway. *Education Inquiry*, 6, 375–399. <https://doi.org/10.3402/edui.v6.27850>
- Elstad, E., Lejonberg, E., & Christophersen, K.-A. (2017). Student evaluation of high-school teaching: Which factors are associated with teachers' perception of the usefulness of being evaluated? *Journal for Educational Research Online*, 9(1), 99–117.
- Ericsson, A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson (Ed.), *The Cambridge handbook of expertise and expert performance* (Vol. 38, pp. 685–705). <https://doi.org/10.1017/CBO9780511816796.038>
- Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Students ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction*, 29, 1–9. <https://doi.org/10.1016/j.learninstruc.2013.07.001>
- Ferguson, R. (2012). Can student surveys measure teaching quality? *Phi Delta Kappan*, 94(3), 24–28. <https://doi.org/10.2307/41763671>
- Fisher, D. L., den Brok, P., & Rickards, T. (2006). Factors influencing students' perceptions of their teachers' interpersonal behaviour: A multilevel analysis. In D. L. Fisher & M. S. Khine (Eds.), *Contemporary approaches to research on learning environments* (pp. 51–74). World Scientific Publishing.
- Gage, N. L. (1960). *Equilibrium theory and behavioral change: An experiment in feedback from pupils to teachers*. Illinois University.
- Gärtner, H. (2014). Effects of student feedback as a method of self-evaluating the quality of teaching. *Studies in Educational Evaluation*, 42, 91–99. <https://doi.org/10.1016/j.stueduc.2014.04.003>
- Gärtner, H., & Brunner, M. (2018). Once good teaching, always good teaching? The differential stability of student perceptions of teaching quality. *Educational Assessment and Educational Accountability*, 30, 159–182. <https://doi.org/10.1007/s11092-018-9277-5>
- Goe, L., Bell, C. A., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. National Comprehensive Center for Teacher Quality.

- Goh, S. C., & Fraser, B. J. (1995, April). *Learning environment and student outcomes in primary mathematics classrooms in Singapore*. Annual meeting of the American Educational Research Association, San Francisco.
- Göllner, R., Fauth, B., Lenske, G., Praetorius, A. K., & Wagner, W. (2020). Do student ratings of classroom management tell us more about teachers or about classroom composition? *Zeitschrift für Pädagogik*, *66*, 156–172.
- Göllner, R., Fauth, B., & Wagner, W. (2021). Student ratings of teaching quality dimensions: Empirical findings and future directions. In W. Rollett, H. J. E. Bijlsma, & S. Röhl (Eds.), *Student feedback in schools. Using student perceptions for improving teaching and teachers* (pp. 111–123). Springer.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hill, H. C., Kapitula, L., & Umland, K. (2011). A validity argument approach to evaluating teacher value-added scores. *American Educational Research Journal*, *48*(3), 794–831. <https://doi.org/10.3102/0002831210387916>
- Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, *64*(4), 349–371. <https://doi.org/10.1037/0021-9010.64.4.349>
- Jones, A. T., & Kirk, C. M. (1990). Gender differences in students' interests in applications of school physics. *Physics Education*, *25*, 308–313.
- Kahmann, K., & Mulder, R. H. (2011). *Feedback in organizations: A review of feedback literature and a framework for future research* (Research Report No. 6). Retrieved from Institute for Educational Science, Universität Regensburg website: https://www.uni-regensburg.de/psychologie-paedagogik-sport/paedagogik-2/medien/kahmann_mulder_2011.pdf
- Kenny, D. A. (2004). PERSON: A general model of interpersonal perception. *Personality and Social Psychology Review*, *8*, 265–280.
- Khalil, N., & Aldridge, J. M. (2019). Assessing students' perceptions of their learning environment in science classes in the United Arab Emirates. *Learning Environments Research*, *22*(3), 365–386. <https://doi.org/10.1007/s10984-019-09279-w>
- Kini, T., & Podolsky, A. (2016). *Does teaching experience increase teacher effectiveness? A review of the research*. Learning Policy Institute.
- Kippers, W. B., Poortman, C. L., Schildkamp, K., & Visscher, A. J. (2018). Data literacy: What do educators learn and struggle with during a data use intervention? *Studies in Educational Evaluation*, *56*, 21–31. <https://doi.org/10.1016/j.stueduc.2017.11.001>
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254–284. <https://doi.org/10.1037/0033-2909.119.2.254>
- Klusmann, U., Kunter, M., Trautwein, U., & Baumert, J. (2006). Lehrerbelastung und Unterrichtsqualität aus der Perspektive von Lehrenden und Lernenden. [Teacher stress and teaching quality from the perspective of teachers and learners]. *Zeitschrift für Pädagogische Psychologie*, *20*(3), 161–173. <https://doi.org/10.1024/1010-0652.20.3.161>
- Krosnick, J. A., & Presser, S. (2010). Questionnaire design. In J. D. Wright & P. V. Marsden (Eds.), *Handbook of survey research* (2nd ed.). Emerald Group.
- Kunst, E. M., van Woerkom, M., & Poell, R. F. (2018). Teachers' goal orientation profiles and participation in professional development activities. *Vocations and Learning*, *11*, 91–111. <https://doi.org/10.1007/s12186-017-9182-y>
- Leigh, A. (2010). Estimating teacher effectiveness from two-year changes in students' test scores. *Economics of Education Review*, *29*(3), 480–488. <https://doi.org/10.1016/j.econedurev.2009.10.010>
- Levy, J., Wubbels, T., Brekelmans, M., & Morganfield, B. (1997). Language and cultural factors in students' perceptions of teacher communication style. *International Journal of Intercultural Relations*, *21*(1), 29–56. [https://doi.org/10.1016/S0147-1767\(96\)00005-3](https://doi.org/10.1016/S0147-1767(96)00005-3)

- Levy, J., Wubbels, T., den Brok, P., & Brekelmans, M. (2003). Students' perceptions of interpersonal aspects of the learning environment. *Learning Environment Research*, 6, 5–36. <https://doi.org/10.1023/A:1022967927037>
- Loeb, S. (2013). *How can value-added measures be used for teacher improvement?* http://www.carnegieknowledge.org/briefs/teacher_improvement/
- London, M., & Smither, J. W. (2002). Feedback orientation, feedback culture, and the longitudinal performance management process. *Human Resource Management Review*, 12, 81–100. [https://doi.org/10.1016/S1053-4822\(01\)00043-2](https://doi.org/10.1016/S1053-4822(01)00043-2)
- Lyden, J. A., Chaney, L. H., Danehower, V. C., & Houston, D. A. (2002). Anchoring, attributions, and self-efficacy: An examination of interactions. *Contemporary Educational Psychology*, 27, 99–117. <https://doi.org/10.1006/ceps.2001.1080>
- Mandinach, E. B., & Gummer, E. S. (2013). Defining data literacy: A report on a convening of experts. *Journal of Educational Research and Policy Studies*, 13(2), 6–28.
- Mandinach, E. B., & Gummer, E. S. (2016). *Data literacy for educators: Making it count in teacher preparation and practice*. Teachers College Press.
- Mandout, L. (2018). Using student feedback to improve teaching. *Educational Action Research*, 26, 755–769. <https://doi.org/10.1080/09650792.2018.1426470>
- Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319–383). Springer.
- Maulana, R., Opendakker, M. C. J. L., den Brok, P., & Bosker, R. J. (2012). Teacher–student interpersonal relationships in Indonesian lower secondary education: Teacher and student perceptions. *Learning Environments Research*, 15(2), 251–271. <https://doi.org/10.1007/s10984-012-9113-7>
- Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26, 237–257. <https://doi.org/10.3102/01623737026003237>
- OECD. (2014). *PISA 2012 results in focus: What 15-year-olds know and what they can do with what they know*. <https://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf>
- Peterson, K. D., Wahlquist, C., & Bone, K. (2000). Student surveys for school teacher evaluation. *Journal of Personnel Evaluation in Education*, 14(2), 135–153.
- Prenger, R., & Schildkamp, K. (2018). Data-based decision making for teacher and student learning: A psychological perspective on the role of the teacher. *Educational Psychology in Practice*, 38, 734–752. <https://doi.org/10.1080/01443410.2018.1426834>
- Prenger, R., Poortman, C. L., & Handelzalts, A. (2017). Factors influencing teachers' professional development in networked professional learning communities. *Teaching and Teacher Education*, 68, 77–90. <https://doi.org/10.1016/j.tate.2017.08.014>
- Rickards, T. (1998). *The relationship of teacher-student interpersonal behaviour with student sex, cultural background and student outcome*. Perth, Australia.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73, 417–458.
- Röhl, S. (2021). Effects of student feedback on teaching and classes: An overview and meta-analysis on intervention studies. In W. Rollett, H. J. E. Bijlsma, & S. Röhl (Eds.), *Student feedback in schools: Using perceptions for the development of teaching and teachers* (pp. 139–156). Springer. https://doi.org/10.1007/978-3-030-75150-0_9
- Röhl, S., & Gärtner, H. (2021). Relevant conditions for teachers' use of student feedback. In W. Rollett, H. J. E. Bijlsma, & S. Röhl (Eds.), *Student feedback in schools: Using perceptions for the development of teaching and teachers* (pp. 157–172). Springer. https://doi.org/10.1007/978-3-030-75150-0_10
- Röhl, S., Bijlsma, H. J. E., & Rollett, W. (2021). The process model of student feedback on teaching (SFT): A theoretical framework and introductory remarks. In W. Rollett, H. J. E. Bijlsma, & S. Röhl (Eds.), *Student feedback in schools: Using perceptions for the development of teaching and teachers* (pp. 1–13). Springer. https://doi.org/10.1007/978-3-030-75150-0_1

- Rowley, J. (2003). Designing student feedback questionnaires. *Quality Assurance in Education*, 11(3), 142–149. <https://doi.org/10.1108/09684880310488454>
- Runhaar, P., Sanders, K., & Yang, H. (2010). Stimulating teachers' reflection and feedback asking: An interplay of self-efficacy, learning goal orientation, and transformational leadership. *Teaching and Teacher Education*, 26, 1154–1161. <https://doi.org/10.1016/j.tate.2010.02.011>
- Schildkamp, K., & Lai, M. K. (2013). Conclusions and a data use framework. In K. Schildkamp & M. K. Lai (Eds.), *Data-based decision making in education: Challenges and opportunities* (pp. 177–191). Springer.
- Schildkamp, K., Lai, M. K., & Earl, L. (2013). *Data-based decision making in education. Challenges and opportunities*. Springer. <https://doi.org/10.1007/978-94-007-4816-3>
- Schildkamp, K., Poortman, C., Luyten, H., & Ebbeler, J. (2017). Factors promoting and hindering databased decision making in schools. *School Effectiveness and School Improvement*, 28(2), 242–258. <https://doi.org/10.1080/09243453.2016.1256901>
- Schweig, J. D. (2016). Moving beyond means: Revealing features of the learning environment by investigating the consensus among student ratings. *Learning Environments Research*, 19(3), 441–462. <https://doi.org/10.1007/s10984-016-9216-7>
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78, 153–189. <https://doi.org/10.3102/0034654307313795>
- Smither, J. W., London, M., & Reilly, R. R. (2005). Does performance improve following multisource feedback?: A theoretical model, meta-analysis, and review of empirical findings. *Personnel Psychology*, 58(1), 33–66. https://doi.org/10.1111/j.1744-6570.2005.514_1.x
- Taylor, S. M., Fisher, C. D., & Ilgen, D. R. (1984). Individuals' reactions to performance feedback in organizations: A control theory perspective. In K. M. Rowland & G. R. Ferris (Eds.), *Research in personnel and human resources management* (pp. 81–124). JAI Press.
- Timm, J. T. (1999). The relationship between culture and cognitive style: A review of the evidence and some reflections for the classroom. *Mid-Western Educational Researcher*, 12(2), 36–44.
- Timmers, C. F., & Veldkamp, B. P. (2011). Attention paid to feedback provided by a computer-based assessment for learning on information literacy. *Computers in Education*, 56(3), 923–930. <https://doi.org/10.1016/j.compedu.2010.11.007>
- Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration*. Ministry of Education. http://www.educationcounts.govt.nz/_data/assets/pdf_file/0017/16901/TPLandDBESentireWeb.pdf
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511819322>
- Veldman, D. J., & Peck, R. F. (1969). Influences on pupil evaluations of student teachers. *Journal in Educational Psychology*, 60(2), 103–108. <https://doi.org/10.1037/h0027085>
- Worthington, A. C. (2002). The impact of student perceptions and characteristics on teaching evaluations: A case study in finance education. *Assessment & Evaluation in Higher Education*, 27(1), 49–64. <https://doi.org/10.1080/02602930120105054>

Hannah J. E. Bijlsma is a researcher at the Dutch School Inspectorate and a primary school teacher (Grade 2). She is interested teaching quality, teacher learning and the improvement of teaching. Her PhD was about the validity and impact of student perceptions of teaching quality.

Sebastian Röhl is postdoctoral researcher in the Institute of Education at Tübingen University (Germany). Among other areas, he conducts research in the fields of teaching development and teacher professionalization through feedback, social networks in inclusive school classes, as well as teachers' religiosity and its impact on professionalism. In addition, he is the director of an in-service professional master's study program for teaching and school development.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 7

Differences in Perceived Instructional Quality of the Same Classrooms with Two Different Classroom Observation Instruments in China: Lessons Learned from Qualitative Analysis of Four Lessons Using TEACH and ICALT



Jieyan Celia Lei, Zhijun Chen, and James Ko 

Abstract Research accumulated has suggested that narrowing instructional quality gaps can improve educational equity and the well-being of children in social and economic backgrounds. Considering that the disparity of instructional quality may affect educational inequality across different regions in China, this study explored how teaching quality varied in 30 lessons primary English classrooms in an economically disadvantaged province in China. This study adopted a mixed-method strategy with quantitative classroom observation data to select four lessons contrastive in teaching quality for subsequent qualitative analysis to explore classroom processes in-depth. Using two internationally validated classroom observation instruments, ICALT and TEACH, added a further dimension to examine how characteristics of instruments might influence perceived instructional quality. Results revealed that while both high-inference instruments were theoretically comparable in distinguishing teaching quality, only ICALT predicted learner engagement. While quantitative instruments could not provide detailed accounts of classroom processes, qualitative accounts of the four lessons could uncover the deep relationships between teacher-student interactions and differences in instructional quality. These

J. C. Lei

Department of Education, Shaoyang University, Shaoyang, China

Department of Education Policy and Leadership, The Education University of Hong Kong, Tai Po, Hong Kong

Z. Chen

Department of Education, University of Bath, Bath, UK

J. Ko (✉)

Department of Education Policy and Leadership, The Education University of Hong Kong, Tai Po, Hong Kong

e-mail: jamesko@eduhk.hk

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_7

findings suggest that conceptually similar instruments may vary in predictive power and that systematic qualitative analysis is indispensable in complementing high-inference instruments to provide an objective teacher evaluation.

Keywords Instructional quality · Classroom observation · Instrument bias

1 Introduction

In the last decade, economically poorer regions worldwide, including inland provinces in China, have received considerable financial support from governmental and non-governmental organisations for building school and teaching and learning facilities equipping to guarantee pupils' schooling. Sammons (2007) identified strong links between school education effectiveness and educational equity and concluded that teacher exerts a substantially more significant effect on children than school, and educational effectiveness varies more at the class level.

Quite a few studies have investigated educational inequalities in China, especially underprivileged areas, from different perspectives such as educational financing (e.g., Li et al., 2007; Tsang & Ding, 2005), gender (e.g., Hannum, 2005; Zeng et al., 2014), poverty (e.g., Heckman & Yi, 2012; Zhang, 2017; Yang et al., 2009), ethnicity (e.g., Hannum et al., 2008, 2015), and urbanisation (e.g., Qian & Smyth, 2008; Yang et al., 2014). Unsurprisingly, educational inequalities in China were found to be narrowed significantly, with the adverse effects primarily mitigated. However, the influence of these factors still exists.

In addition to non-classroom observation factors, classroom teaching quality directly impacts students' learning effectiveness. Given the significant role of classroom teaching practices in greater educational equity (Sammons, 2007), a research gap lies in the lack of lesson observation evidence on the quality of classroom teaching exploration in an underprivileged area in China. Furthermore, the rapid development of China society in recent years makes studies easily and quickly outdated. Lack of timely updated research prevents audiences' knowledge of the education situation from keeping pace with reality. This study explored educational inequality at the classroom teaching level from a teaching effectiveness perspective in an under-advantaged province in China. Using two classroom observation instruments, ICALT (Van de Grift, 2007) and TEACH (World Bank, 2019), we explored the instructional quality gaps between example lessons and how the perceived instructions differed in learning and teaching interactions.

2 Literature Review

2.1 *Teaching Quality in Developing Countries and Underdeveloped Regions*

Factors affecting students' outcomes at the classroom level have received more attention than factors at the school level in educational effectiveness research (Muijs et al., 2014). Knowledge in effective teaching practice at the classroom level is crucial for enhancing teacher capability to develop agile differentiated instruction strategies for diverse learners' needs (Edwards et al., 2006). Although strenuous efforts have been made to probe into teaching quality in classrooms, studies between developed and developing countries are insufficient. The Organisation for Economic Cooperation and Development's PISA 2018 project (OECD, 2019), which evaluated the academic performance of junior secondary students worldwide, involved only two developing countries/regions among the 30 participating countries/regions.

We generally lack knowledge in classroom-level teaching quality in developing countries/regions except for a few noticeable empirical studies. For example, Chiangkul (2016) claimed that insufficient capability in the knowledge and teaching skills of the younger Thai teachers was evident in the *Trends International Mathematics and Science Study* (TIMSS) 2015. In South Africa and Botswana, teachers were found to lack knowledge about combining practical pedagogical skills with subject content (Sapire & Sorto, 2012). In rural Guatemala, Marshall and Sorto (2012) found that teaching practice in mathematics classrooms adopted less complex pedagogical skills than developed countries like Japan, America and Germany. Similarly, teaching quality in China varies province by province, and inland provinces have disadvantages noticeably in recruiting talented teachers. Moreover, the teaching capability of rural schoolteachers was generally lower than that of urban teachers, resulting in a remarkable gap between rural and urban schools in West China (Wang & Li, 2009). Thus, understanding teaching effectiveness in rural regions of economically disadvantaged provinces in China would contribute to strategies to promote educational quality and equity for children in the regions in the future.

2.2 *Classroom Observation and Comparison of Instruments*

Studies of student academic outcomes significantly contribute to classroom effectiveness, but the specific processes are not articulated (Pianta et al., 2008). The invention of classroom observation instruments provides a powerful approach for probing into classroom reality. It is seen as a more just form of data collection to examine teachers' behaviours (Pianta et al., 2008). Classroom observation used to be limited to teacher appraisal, lesson evaluation, professional development of novice teachers, identifications of expert teachers from experienced teachers, but it has

become popular with the interest in the classroom level teaching process in research increased (Wragg, 2013). Systematic classroom observation allows teachers to compare specific predetermined and agreed categories of behaviour and practice, which originated in teacher effectiveness research (Muijs & Reynolds, 2005).

Lesson videos of classroom teaching practice could be another observation form that provides researchers with a window to explore what happens in classrooms (Sapire & Sorto, 2012). For teaching analysis, video data was first used in the TIMSS 1995 video study by Stigler et al. (1999). Video recordings allow raters to slow down, pause, replay and re-interpret teaching practice, and capture complex teaching paths (Erickson, 2011; Jacobs et al., 1999; Klette, 2009). Furthermore, recorded teaching practice makes visual representation possible for researchers to capture anticipated details of classrooms that may escape their gaze (Lesh & Lehrer, 2000; Tee et al., 2018).

A few observation instruments were developed to evaluate teachers' actual teaching processes and their contribution to student achievements. For exploring the generic pedagogic capability of teachers, these observational tools include *the Framework for Teaching* (Danielson, 1996), *the International System for Teacher Observation and Feedback* (Teddlie et al., 2006), *the International Comparative Analysis of Learning and Teaching* (ICALT) (Van de Grift, 2007), *the Classroom Assessment Scoring System* (CLASS) (Pianta et al., 2008), and the TEACH (World Bank, 2019). Some assess specific competencies, such as classroom talk (Mercer, 2010) and project-based learning (Stearns et al., 2012). Instruments for subject-specific pedagogies are available to researchers as well, such as English reading (Gersten et al., 2005), mathematical instruction (Schoenfeld, 2013) and historical contextualisation (Huijgen et al., 2017).

For instrument application, scholars compared different instruments for STEM classrooms in post-secondary education (Anwar & Menekse, 2021), mathematics and science classrooms in secondary education (Boston et al., 2015; Marshall et al., 2011) and preservice teacher internships (Caughlan & Jiang, 2014; Henry et al., 2009). However, no instruments comparison study based on English as a second language classrooms in primary education was found, which could contribute to essential education quality improvement in developing countries.

In the present study that compared ICALT and TEACH, we identified two issues in our careful comparisons of the two instruments. First, theoretically speaking, the two instruments are conceptually similar. The teaching behaviours under the *Classroom Culture* domain of TEACH are conceptually similar to the behavioural indicators of the *Safe and Stimulating Learning Climate* and *Efficient Organisation* domains of ICALT (Van de Grift, 2007). Similarly, *the Socioemotional Skills* domain of TEACH is conceptually comparable to the *Intensive and Activating Teaching* domain of ICALT. The *Instruction* domain of TEACH is similar to ICALT's *Clear and Structured Instructions*, *Adjusting Instructions* and *Learner Processing to Inter-Learner Differences* and *Teaching Learning Strategies* domains.

The inspectors initially developed ICALT to study primary classrooms in England and the Netherlands. The ICALT was then used as a research tool to compare teaching practices in developed and developing countries (Maulana et al.,

2021). In contrast, TEACH was developed as a system diagnostic and monitoring tool of teaching practices at a primary school level to foster professional development in low- and middle-income countries (Molina et al., 2018). Thus, the difference in scale development would be, theoretically and methodologically, critical if TEACH is more suitable for developing regions or countries than ICALT. For example, it is unlikely that catering for learner diversity is considered essential in developing countries where access to free education is challenging. Maulana et al. (2021) have shown that teaching behaviours associated with differentiation could be country-specific rather than universal.

Second, it is less difficult to conduct classroom observation with TEACH in practice than ICALT. ICALT was designed to observe whether teachers adjust teaching according to the level of students, but ICALT also emphasises stimulating students with weak learning abilities to build self-confidence. This teaching behaviour reflects a higher teaching skill of teachers. Kyriakides et al. (2009) found that teacher behaviours varied distinctively in difficulty levels, and it is not uncommon that teachers cannot master some advanced teaching skills even after professional training. Similarly, Ko et al. (2015) found that while teachers in Guangzhou were found performing better than Hong Kong teachers in many aspects of perceived teaching quality, Hong Kong teachers did better in catering for learner diversity because Hong Kong has practised an inclusive education policy for nearly two decades.

2.3 *Qualitative In-Depth Lesson Analysis from a Dialogic Teaching Perspective*

Apart from the dominant quantitative teacher effectiveness research, a consistently growing body of research investigated learning and teaching from a qualitative perspective on dialogic teaching in the last decades (Howe & Mercer, 2017; Vrikki et al., 2019) with regarding dialogic teaching as vital to student learning outcome (Alexander, 2006; Howe et al., 2019). Alexander (2008) proposed dialogic teaching as a learning process that promotes students to develop their higher-order thinking through reasoning, discussing, arguing, and explaining. Dialogic teaching is believed to have two main types, teacher-student interaction and student-student interaction (Howe & Abedin, 2013), with five core principles: collective, reciprocal, supportive, cumulative and purposeful (Alexander, 2008).

Hennessy and his team (2016) introduced a coding approach with developed *Scheme for Education Dialogue Analysis* (SEDA) to conduct qualitative in-depth lesson analysis for characterising and analysing classroom dialogues. It is considered a practical approach to evaluate how high-quality interaction is productive for learning (Hennessy et al., 2020), and has become quite prevalent in recent years (Song et al., 2019). For example, Shi et al. (2021), informed by SEDA's condensed version, the Cambridge Dialogue Analysis Scheme (CDAS) (Vrikki et al., 2019), successfully modified SEDA to make it more suitable for their data set.

3 Research Questions

Based on the above background and consideration, the objective of this study is to answer the following research questions:

1. How were teaching practices rated using different classroom instruments (i.e., ICALT and TEACH) in the same lessons?
 - (a) In what aspects did the ratings look similar based on the two observation instruments?
 - (b) How did the rating show more variations based on the two observation instruments?
2. To what extent the above differences could be identified in an in-depth qualitative analysis of four purposively selected lessons?

4 Method

This study adopted a subsequent quantitative-qualitative research strategy to probe into the link and differences between two instructional quality assessment instruments, the TEACH and the ICALT. This research used the classroom observation strategy to explore teachers' teaching quality and teacher-student interactions.

4.1 Samples

This study involved 20 primary schools in an underprivileged province in China in two different districts (one city/urban and one county/rural). Among these twenty schools, eleven schools were from the rural area, and nine were from the urban area. Thirty English teachers (one lesson per teacher) randomly selected from the sample schools participated in this study. The data collection was conducted with a third party that targeted primary school teachers whose teaching experience was more than two years and less than eight years. Hence, we controlled the teaching experience of participants by excluding teachers with less than two years or more than eight years.

Thirty lessons (one lesson per teacher) were recorded and observed by a well-trained rater with instruments to obtain quantitative data. Then, four lessons were selected for in-depth qualitative analysis.

4.2 Instruments

Classroom observation instruments are often assumed to study similar teaching characteristics, so they are expected to be comparable (Ko, 2010). ICALT (Van de Grift, 2007) and TEACH (World Bank, 2019) are two internationally validated classroom observation instruments on generic teaching behaviours. Analysis of this study focuses on high-inference indicators of these two instruments.

4.2.1 ICALT

ICALT instrument (Van de Grift, 2007) assesses classroom teaching behaviours divided into three parts. The core part has 32 behavioural indicators to be evaluated on a four-point scale to determine the relative strengths and effectiveness of a teaching behaviour (i.e., 1 = mostly weak; 2 = more often weak than strong; 3 = more often strong than weak; 4 = mostly strong). Four to ten behavioural indicators are grouped in one of the six primary domains in the instrument: *Safe and Stimulating Learning Climate*, *Efficient Organisation*, *Clarity and Structure of Instruction*, *Intensive and Activating Teaching*, *Adjusting Instructions and Learner Processing to Inter-Learner Differences* groups, and *Teaching Learning Strategies*. The second part comprises 115 observable teaching behaviours, with 3–10 matching a behavioural indicator in the core part. For example, ‘*The teacher lets learners finish their sentences,*’ ‘*The teacher listens to what learners have to say,*’ and ‘*The teacher does not make role stereotyping remarks*’ are corresponding teaching behaviours for the first indicator, ‘*The teacher shows respect for learners in his/her behaviour and language*’. Before giving a score for the behavioural teaching indicators, a rater should determine whether the observed behaviours are observed during the lesson. Whenever a teaching behaviour is observed, it should be scored 1; or a zero should be given if it is not observed. This part of ICALT has made the instrument quite different from many other instruments (e.g., the *Classroom Assessment Scoring System* by Pianta et al., 2008; Pianta & Hamre, 2009) because a rater is expected to judge the effectiveness of a teaching indicator on the grounds of a set of observed teaching behaviours. The last part of ICALT includes three behavioural indicators for learner engagement and ten associated learning behaviours, evaluated in 4-point and 2-point respectively.

4.2.2 TEACH

TEACH was a validated classroom observation tool developed by the World Bank (2019), applicable for Grade 1–6 classrooms in primary schools. It aimed to promote teaching quality improvement in under-advantaged nations. Raters of this instrument showed high inter-rater reliability (Molina et al., 2018). This instrument offers a unique window into some seldom investigated but weighty domains of class

level teaching and learning experiences. The *Time on Task* component requires observers to record in three ‘snapshots’ of 1–10 seconds whether teachers provide most students with learning activities and how many students are on task. *Classroom Culture, Instruction, and Socioemotional Skills* are the three domains of the *Quality of Teaching Practice* component, followed by nine corresponding indicators that point to 28 teaching behaviours. Based on observation reality, observers rate each behaviour item with a three-level scale, ‘high’, ‘medium’ and ‘low’, equal to ‘definitely having this behaviour’, ‘somewhat having this behaviour’ and ‘only having opposite behaviour’ respectively. It should be noted that four behaviour items can be marked as ‘N/A’ if they do not occur in the classroom. By matching its corresponding behavioural ratings, each indicator is scored with a five-point scale, ranging from 1 to 5 (‘1’ is the lowest and ‘5’ is the highest).

4.2.3 Comparison of ICALT and TEACH

Through careful comparisons at the level of behavioural indicators, it was found that the teaching behaviours under the *Classroom Culture* domain of TEACH correspond to the behavioural indicators of the *Safe and Stimulating Learning Climate* and *Efficient Organisation* domains of ICALT (Van de Grift, 2007). Similarly, the *Socioemotional Skills* domain of TEACH corresponds to the *Intensive and Activating Teaching* domain of ICALT. The *Instruction* domain of TEACH corresponds to ICALT’s *Clear and Structured Instructions, Adjusting Instructions and Learner Processing to Inter-Learner Differences* and *Teaching Learning Strategies* domains. It is less difficult to conduct classroom observation with TEACH than ICALT. As mentioned earlier, while ICALT and TEACH could be used to observe whether teachers adjust teaching according to student abilities, the *Adjusting Instructions and Learner Processing to Inter-Learner Differences* domain in ICALT also emphasises stimulating students with weak learning abilities to build self-confidence. This domain reflects a higher level of teaching skills of teachers.

However, as a specific classroom observation instrument for teacher evaluation in primary schools in underdeveloped countries, TEACH is a better choice for in-depth qualitative analysis on dialogic teaching with its official training manual (World Bank, 2019), providing clear definitions on teaching behaviour items and detailed guidance for observer training. All teaching behaviour indicators in TEACH have unified official inspection standards, ensuring the reliability of coding scheme building and the in-depth qualitative dialogue analysis process and results. Accordingly, a new qualitative coding scheme, *TEACH Tool for Lesson Analysis* (TTLA), was developed based on the TEACH manual and partially summarised in Table 7.1.

Table 7.1 TEACH tool for lesson analysis (TLA)—A qualitative coding scheme based on the TEACH framework

Area code	Area name	Description
C	CLASSROOM CULTURE	<p>The teacher creates a culture that is conducive to learning. The focus here is not on the teacher correcting students' negative behaviours but rather the extent to which the teacher creates:</p> <ul style="list-style-type: none"> (i) a supportive learning environment by treating all students respectfully, consistently using positive language, responding to students' needs, and both challenging gender stereotypes and not exhibiting gender bias in the classroom; and (ii) positive behavioural expectations by setting clear behavioural expectations, acknowledging positive student behaviour, and effectively redirecting misbehaviour.
I	INSTRUCTION	<p>The teacher instructs to deepen student understanding and encourage critical thinking and analysis. The focus here is not on content-specific methods of instruction, but rather the extent to which the teacher:</p> <ul style="list-style-type: none"> (i) facilitates the lesson by explicitly articulating lesson objectives that are aligned to the learning activity, clearly explaining content, and connecting the learning activity to other content knowledge or students' daily lives, and by modelling the learning activity through enacting or thinking aloud; (ii) does not simply move from one topic to the next but checks for understanding by using questions, prompts, or other strategies to determine students' level of understanding, by monitoring students during group and independent work, and by adjusting his/her teaching to the level of students; (iii) gives feedback by providing specific comments or prompts to help clarify students' misunderstandings or identify their successes; and (iv) encourages students to think critically by asking open-ended questions and providing them with thinking tasks requiring them to analyse content actively. Students exhibit critical thinking ability by asking open-ended questions or performing thinking tasks.
S	SOCIOEMOTIONAL SKILLS	<p>The teacher fosters socio-emotional skills that encourage students to succeed inside and outside the classroom. To develop students' social and emotional skills, the teacher:</p> <ul style="list-style-type: none"> (i) instils autonomy by allowing students to make choices and take on meaningful roles in the classroom. Students exhibit their autonomy by volunteering to participate in classroom activities; (ii) promotes perseverance by acknowledging students' efforts, rather than focusing solely on their intelligence or natural abilities, by having a positive attitude toward students' challenges by framing failure and frustrations as part of the learning process, and by encouraging students to set short- and long-term goals; and (iii) fosters social and collaborative skills by encouraging peer interaction and promoting interpersonal skills, such as perspective-taking, empathising, emotion regulation, and social problem-solving. Students exhibit social and collaborative skills by collaborating through peer interaction.

(continued)

Table 7.1 (continued)

	Element name	Description
CLASSROOM CULTURE		
C1	SUPPORTIVE LEARNING ENVIRONMENT	The teacher creates a supportive learning environment. The teacher creates a classroom environment where students feel emotionally safe and supported. Moreover, all students feel welcome, as the teacher treats all students respectfully.
C2	POSITIVE BEHAVIORAL EXPECTATIONS	The teacher promotes positive behaviour in the classroom. The teacher promotes positive behaviour by acknowledging students' behaviour that meets or exceeds expectations. Moreover, the teacher sets clear behavioural expectations for different lesson parts.
INSTRUCTION		
I3	LESSON FACILITATION	The teacher facilitates the lesson to promote comprehension. The teacher facilitates the lesson to promote comprehension by explicitly articulating the objectives, providing clear explanations of concepts, and connecting the lesson with other content knowledge or students' experiences.
I4	CHECKS FOR UNDERSTANDING	The teacher checks the understanding of most students. The teacher checks for understanding to ensure most students comprehend the lesson content. Moreover, the teacher adjusts the pace of the lesson to provide students with additional learning opportunities.
I5	FEEDBACK	The teacher provides feedback to deepen student understanding. The teacher provides specific comments or prompts to help identify misunderstandings, understand successes, and guide thought processes to promote learning.
I6	CRITICAL THINKING	The teacher builds students' critical thinking skills. The teacher builds students' critical thinking skills by actively encouraging them to analyse content.
SOCIOEMOTIONAL SKILLS		
S7	AUTONOMY	The teacher allows students to make choices and encourages students to participate in the classroom. The teacher provides students with opportunities to make choices and take on meaningful roles in the classroom. Students use these opportunities by volunteering to take on roles and expressing their ideas and opinions throughout the lesson.
S8	PERSEVERANCE	The teacher promotes students' efforts, has a positive attitude toward challenges, and encourages goal setting. The teacher promotes students' efforts toward mastering new skills or concepts instead of focusing solely on results, intelligence, or natural abilities. In addition, the teacher has a positive attitude toward challenges, framing failure and frustrations as valuable parts of the learning process. The teacher also encourages students to set short- and/or long-term goals.
S9	SOCIAL & COLLABORATIVE SKILLS	The teacher fosters a collaborative classroom environment. The teacher encourages students' collaboration and promotes students' interpersonal skills. Students respond to the teacher's efforts by collaborating in the classroom, creating an environment free from physical or emotional hostility.

(continued)

Table 7.1 (continued)

TEACH behaviours	Keywords	Description
C1a	Respect students	The teacher treats all students respectfully. For example, The teacher uses students' names, says "please" and "thank you," or shows some other culturally relevant sign of respect.
C1b	Positive language	The teacher consistently uses positive language in his/her communication with students. For example, The teacher consistently uses encouraging phrases such as "Great job!" when students show their work to him/her, or "You can do this!", or "You are such a talented group of children."
C2a	Behaviour expectation	The teacher sets clear behavioural expectations for classroom tasks and/or activities throughout the lesson. For example, upon introducing a group activity to the class, the teacher explicitly states the expected behaviour of students in the group. His expectations may include, "Use a quiet indoor voice" or "Take turns speaking." Alternatively, the teacher is not observed setting clear behavioural expectations, but students are well-behaved ⁵ throughout the lesson.
I3c	Connect lesson to Ss' life	The teacher meaningfully connects the lesson to other content knowledge or students' daily lives. For example: When teaching a class on fractions, the teacher relates the content to students' experiences by asking, "Who has had to slice a birthday cake? How did you make sure there were enough slices for everyone? Learning about fractions can help us divide a cake between people." The teacher also connects the lesson to a previous lesson on halves by saying, "Remember yesterday when we learned about halves? We learned that when we cut a cake in half, we can share it equally between 2 people. Today, we will learn how to divide the cake into fourths to share the cake. When we formed halves, we made sure we had two halves of identical size. The same thing is true when we are forming fourths: we have to make sure to keep slices of the same size." The connection between the current lesson and other content knowledge and/or students' daily lives is unmistakable.
I3d	Modle by enacting/ thinking aloud	The teacher completely models the learning activity by enacting all parts of the procedure OR by enacting the procedure AND thinking aloud. For example, The teacher demonstrates different ways to solve a math problem (enactment of a procedure) and while doing so, s/ he says what s/he is thinking at each step of the equation (think aloud). If students calculate the areas of their desks, the teacher demonstrates each step in the process (full enactment of a procedure).
I4a	Determine Ss' understanding	The teacher uses questions, prompts, or other strategies that effectively determine most students' level of understanding. For example, The teacher says, "Please put your thumb up if you agree or down if you disagree with this statement: Equilateral triangles have equal angles." The teacher also asks students to demonstrate their knowledge by having all students share their answers, e.g., asking each student to read out the sentence s/he wrote using past tense verbs.

(continued)

Table 7.1 (continued)

TEACH behaviours	Keywords	Description
I6a	Ask open-end question	The teacher asks students three or more open-ended questions, AND at least 1 of them builds upon student responses by asking students to justify their reasoning, further explain, or clarify their ideas. For example, The teacher asks, “How do you think the main characters in the story would prepare for the competition?” After a student responds, the teacher then follows up by asking, “What facts or ideas make you think that?” Then s/he asks another student, “What do you think happens next?” In a math class, the teacher asks, “How do you know -2 is greater than -6 ?” After the student responds, the teacher follows up by asking, “What would happen if the numbers were positive?” Later in the lesson, the teacher asks, “How do you use the number line to determine if -8 or -4 is greater?”
I6c	Ss ask open-end Questions/ perform thinking tasks	Students ask open-ended questions. For example, after working on subtraction problems, a student asks, “Why does $6-9$ equal a negative number?” Alternatively, they perform substantial thinking tasks.
S7b	Opportunities for Ss to take on roles	The teacher provides students with opportunities to take on meaningful roles in the classroom, in which they are responsible for parts of a learning activity. For example, The teacher allows a student to solve an equation on the blackboard and explain how s/he tackles the main challenges of a problem.
S7c	Ss volunteer to participate	Most students volunteer to participate by expressing their ideas and taking on roles. For example: When the teacher asks a question, many students put their hands up to share their answers. The students could also volunteer without the teacher asking (e.g., a student offers to share a related experience when explaining a concept).

4.3 *Raters*

The first author served as a research assistant in a commissioned impact study in which she collected all videos while she observed, recorded and rated with TEACH all the lessons onsite. Then, she reviewed the lesson videos with ICALT again within a month. The rater held a master’s degree with considerable lesson observation experience after taking TEACH and ICALT training workshops. The first author evaluated the same lesson videos with two instruments in the workshops and conducted a comparison and discussion afterwards. Then the raters launched the second and third rounds of lesson video evaluation practice. An additional rater was employed to ensure better consistency on inter-rater reliability concerns. The rater informed teachers only one night before the observation to prevent teachers from preparing perfect teaching in advance. All 30 classrooms were recorded with a camera to enable later transcripts on teaching practice and in-depth coding of teaching behaviours.

4.4 Data Collection

4.4.1 Quantitative Rating

A total of thirty English lessons were observed. Quantitative analysis was conducted with SPSS 20 to compare the perceived instructional quality of the same classrooms in different aspects of classroom observation instruments, TEACH and ICALT and determine which instrument could better predict student engagement. As Z-scores averages were provided in the official manual of TEACH (World Bank, 2019), selecting lessons for comparison based on those averages would provide objective ground beyond the present study. Two ‘weak’ lessons (Lesson 1, $z = -1.52$; Lesson 2, $z = -0.96$) and two ‘strong’ lessons (Lesson 3, $z = 1.24$; Lesson 4, $z = 2.62$) were eventually selected for in-depth qualitative analyses to explore variations in the evaluations of teaching quality with different instruments (see Table 7.1).

4.4.2 Qualitative Coding

In-depth qualitative analyses were performed based on the teaching behaviour definitions in the TEACH manual for better validity. TTLA was employed to code the teaching behaviours of the four selected four lessons. Teaching activities and interactions between teachers and students of each sample lesson illustrated teaching practices more specifically than quantitative ratings.

5 Results

5.1 Quantitative Analyses of All Lessons

All TEACH and ICALT factors were standardised for quantitative analyses because the scales used were different in the two instruments. Due to the small sample sizes, only one regression model was tested using SPSS 20.0 to predict learner engagement in ICALT using the overall scores of both TEACH and ICALT.

Table 7.2 presents the mean, standard deviation, and reliability (alpha and omega) of factors in two instruments. We include both McDonald’s Omega (McDonald, 2013) and Cronbach’s alpha (1951), as the former is considered more suitable regardless of the number of items within a factor. The results indicated that the two values do not show much difference. It also demonstrates the descriptive statistics of the overall scores and good item consistencies of all nine items in TEACH ($\alpha = 0.82$) and 32 items in ICALT ($\alpha = 0.932$). Due to a limited number of items in each TEACH factor, there is a low internal consistency level for *Socioemotional Skills* ($\alpha = 0.483$). In ICALT, the *Adjusting Instructions and Learner Processing to Inter-Learner Differences* domain ($\alpha = 0.361$) and *Teaching Learning Strategies* domain ($\alpha = 0.599$) also show low reliabilities.

Table 7.2 Mean, standard deviation and reliability of factors in TEACH and ICALT

	Mean	Std. deviation	Alpha	Omega	N of items
<i>Classroom culture</i>	3.433	0.640	0.478		2
<i>Instruction</i>	2.475	0.407	0.700	0.751	4
<i>Socioemotional skills</i>	2.578	0.446	0.478	0.483	3
TEACH average	2.722	0.403	0.768	0.82	9
Safe and stimulating learning climate	2.433	0.565	0.844	0.849	4
Efficient organisation	3.000	0.572	0.691	0.705	4
Clear and structured instructions	2.624	0.456	0.848	0.87	7
Intensive and activating teaching	1.976	0.379	0.658	0.691	7
Adjusting instructions and learner processing to inter-learner differences	1.208	0.198	0.206	0.361	4
Teaching learning strategies	1.106	0.183	0.490	0.599	6
ICALT average	2.044	0.306	0.916	0.932	32
Learner engagement	2.344	0.750	0.929	0.931	3

Spearman rho's correlation coefficients between TEACH and ICALT factors are presented in Table 7.3. There are strong positive correlations between three TEACH factors, while the ICALT domain *Adjusting Instructions and Learner Processing to Inter-learner Differences* does not significantly correlate with other ICALT domains. *Learner engagement* was significantly correlated with most factors in both TEACH and ICALT, except for the *Adjusting Instructions and Learner Processing to Inter-Learner Differences* domain in ICALT.

With the limitation of the participant number, only one regression model with the overall scores of TEACH and ICALT in the prediction of learner engagement could be conducted (see Table 7.4). Results show that only the ICALT score could significantly predict learner engagement, $F(2, 27) = 29.92$, $p < .00$, $R^2 = 0.83$.

5.2 Comparisons of ICALT and TEACH Results of the Selected Four Lessons

As shown in Table 7.5, the individual and overall aspects of LESSON 1 and LESSON 2 were relatively weak with lower means, while LESSON 3 and LESSON 4 were high-quality lessons. The standard deviations of the ICALT averages (Table 7.5) were observably lower than that of TEACH, indicating that variations in ratings were more considerable if TEACH was used for observation.

At the domain level, LESSON 1 has a much lower mean in the *Instruction* domain ($M = 1.75$) but a little higher means in the *Classroom Culture* ($M = 2.5$) and *Socioemotional Skills* ($M = 2.33$) domains than those of LESSON 2 ($M = 2.75$, 2.0 , 2.0 respectively) in the TEACH results. However, the ICALT results show LESSON 1 scored much higher means in the *Safe and Stimulating Learning Climate* domain

Table 7.3 Correlations (Spearman rho) between TEACH (1–3) and ICALT factors (4–9)

	1	2	3	4	5	6	7	8	9	10
1. Classroom culture (TEACH)	1									
2. Instruction (TEACH)	.548**	1								
3. Socioemotional skills (TEACH)	.630**	.604**	1							
4. Safe and stimulating learning climate (ICALT)	.668**	.602**	.672**	1						
5. Efficient organisation (ICALT)	.764**	.543**	.645**	.735**	1					
6. Clear and structured instructions (ICALT)	.440*	.706**	.578**	.700**	.539**	1				
7. Intensive and activating teaching (ICALT)	.454*	.686**	.655**	.697**	.438*	.833**	1			
8. Adjusting instructions and learner processing (ICALT)	0.106	0.153	-0.063	0.289	0.170	0.068	0.197	1		
9. Teaching learning strategies (ICALT)	.417*	.633**	.522**	.536**	0.354	.698**	.771**	0.180	1	
10. Learner engagement (ICALT)	.391*	.608**	.618**	.708**	.540**	.778**	.768**	0.099	.623**	1

** indicates $p < 0.01$; * indicates $p < 0.05$

($M = 2.25$) and a little higher in the *Intensive and Activating Teaching* domain ($M = 1.57$), and a little lower mean in *Clear and Structured Instructions* domain ($M = 2.29$) than LESSON 2. It is worth noting that the ICALT rankings of these two less effective lessons are higher than those of TEACH. Interestingly, LESSON 1 ranks the last in TEACH but the 22nd out of 30 in ICALT. LESSON 2 ranks higher than LESSON 1 in TEACH (28th) but higher in ICALT (26th).

Regarding the two more effective lessons, means of LESSON 3 in the *Instruction* ($M = 3.0$) and *Socioemotional Skills* ($M = 3.0$) domains are significantly lower than LESSON 4 ($M = 3.75, 3.67$ respectively) in TEACH. In contrast, for ICALT, means

Table 7.4 Linear regression model using learner engagement in ICALT as a dependent variable

	B	p-value
(Constant)	0.00	1.000
<i>Classroom culture</i>	0.349	0.059
<i>Instruction</i>	0.641	<0.001
<i>Socioemotional skills</i>	0.667	<0.001
Safe and stimulating learning climate	0.728	<0.001
Efficient organisation	0.502	0.005
Clear and structured instructions	0.796	<0.001
Intensive and activating teaching	0.78	<0.001
Adjusting instructions and learner processing to inter-learner differences	0.1	0.599
Teaching learning strategies	0.577	<0.001
ICALT_average	0.829	<0.001
TEACH_average	0.657	<0.001

for LESSON 3 were lower in the *Safe and Stimulating Learning Climate* ($M = 3.0$), *Intensive and Activating Teaching* ($M = 2.29$), *Adjusting Instructions and Learner Processing to Inter-Learner Differences* ($M = 1.0$), and *Teaching Learning Strategies* ($M = 1.0$) domains than those for LESSON 4 ($M = 3.5, 2.71, 1.5, 1.67$ respectively). LESSON 3 were rated better in two ICALT domains, *Efficient Organisation* ($M = 3.75$) and *Clear and Structured Instructions* ($M = 3.57$), than LESSON 4 ($M = 3.5, 2.71$ respectively). Additionally, the ranking of two high-quality lessons of TEACH was a little higher than that of ICALT. LESSON 3 ranks 3rd in TEACH but 5th in ICALT, and LESSON 4 ranks 1st in TEACH and 2nd in ICALT.

5.3 Qualitative Characteristics of Teacher-Student Interactions

Two low-quality lessons (LESSONS 1 & 2) and two high-quality lessons (LESSONS 3 & 4) were selected as above mentioned. Four lessons were transcribed verbatim and coded with non-verbal communication captured by two coders. Coders coded these lessons with the TTLA framework outlined in Table 7.1. Teaching behaviours reflected in dialogue content are coded with corresponding codes. Multiple coding appears when more than one behaviour is reflected.

The performances of two low-quality lessons (LESSONS 1 & 2) were unsatisfactory in the teacher-student interaction. Table 7.6 shows the learning activity Reading Sentences of LESSON 1. The teacher performed good at providing students with opportunities to play a role in the classroom (S7b) and promoted students' voluntary behaviours (S7c). Nevertheless, students were not clear with the learning activity behaviour expectation since the teacher did not explain it before the learning activity. When the teacher said, 'partner A partner B', all students were confused and silent (Line 2). They had no idea what the teacher expected them to do until she asked who wanted to be Partner A in English and Chinese.

Table 7.5 Comparisons of four lessons in TEACH and ICALT scores

	<i>TEACH_</i> <i>Classroom</i> <i>culture</i>	<i>TEACH_</i> <i>Instruction</i>	<i>TEACH_</i> <i>Socio-</i> <i>emotional</i> <i>skills</i>	<i>ICALT_</i> Safe and stimulating learning climate	<i>ICALT_</i> Efficient organisation	<i>ICALT_</i> Clear and structured instructions	<i>ICALT_</i> Intensive and activating teaching	<i>ICALT_</i> Adjusting instructions and learner processing to inter-learner differences	<i>ICALT_</i> Teaching learning strategies	<i>ICALT_</i> Learner engagement	Teach average	<i>ICALT_</i> average	Z _{TEACH}	TEACH Rank	<i>ICALT_</i> Rank
Lesson 1	2.50	1.75	2.33	2.25	2.50	2.29	1.57	1.25	1.00	1.67	2.11	1.78	-1.52	30	22
Lesson 2	2.00	2.75	2.00	1.25	2.50	2.57	1.43	1.25	1.00	1.67	2.33	1.69	-0.96	28	26
Lesson 3	4.00	3.00	3.00	3.00	3.75	3.57	2.29	1.00	1.00	3.67	3.22	2.44	1.24	3	5
Lesson 4	4.00	3.75	3.67	3.50	3.50	2.71	2.71	1.50	1.67	3.67	3.78	2.56	2.62	1	2
4 lesson_ mean	3.13	2.81	2.75	2.50	3.06	2.79	2.00	1.25	1.17	2.67	2.86	2.12			
4 lesson_ SD	1.03	0.83	0.74	0.98	0.66	0.55	0.61	0.20	0.33	1.15	0.78	0.45			
Overall_ Mean	3.43	2.48	2.58	2.43	3.00	2.62	1.98	1.21	1.11	2.34	2.72	2.04			
Overall_ SD	0.64	0.41	0.45	0.56	0.57	0.46	0.38	0.20	0.18	0.75	0.40	0.31			

Table 7.6 Lesson 1 Reading sentences

Agent	Line	Reading sentences	TA			
			Code1	Code2	Code3	Code4
Teacher	1	Partner A. Partner B.	S7b			
Students	2	(Silence).	NA			
Teacher	3	Who is partner A? 谁 (who) 是 (is)A?	S7b			
Student A	4	(Several students raised their hands. The teacher invited a student with a gesture).	S7b	S7c		
Teacher	5	Partner B. 谁 (who)?	S7b			
Student B	6	(Several students raised their hands. The teacher invited a student with a gesture).	S7b	S7c		
Teacher	7	好 (Good). A.	S7b			
Student A	8	They had some food.	S7b			
Student B	9	They had some drinks.	S7b			
Teacher	10	Again.	S7b			
Student A	11	They had some food.	S7b			
Student B	12	They had some drinks.	S7b			
Teacher	13	Now, (you are) group A. Now, (you are) group B, OK?	S7b			
Students	14	OK.	S7b			
Teacher	15	OK. Again.	S7b			
Student A	16	They had some food.	S7b			
Students	17	They had some food.	S7b			
Student B	18	They had some drinks.	S7b			
Students	19	They had some drinks.	S7b			
Teacher	20	OK. Set down.	C2a			

The situation in LESSON 2 (Table 7.7) was also difficult. The teacher in LESSON 2 performed poorly in respecting students. The teacher even taunted the students (line 7: *Aren't you full? Can't the brain think?* [means *You are a fool* in Chinese culture]). On the bright side, the teacher offered students opportunities to play a role in the classroom (9 lines out of 10 lines of teacher talk were coded with S7b) by asking questions to check students' level of understanding (I4a). However, he did not tell students what they could refer to and where the references were in advance, so it was hard to follow him. Students responded to the teachers' questions with silence (Line 4, Line 6, Line 11, Line 13), making the lesson challenging to move on.

As one of the high-quality lessons, LESSON 3 led the students to review the words learned before (Table 7.8). First, the teacher explained the expected behaviours of the learning activity and demonstrated how to carry out the activity in detail, and even conducted simulation (Line 7, C2a, I3d; Line 9, I3d; Line 11, C2a; Line 13, C2a). In this activity, the teacher attached great importance to students' mastery of learning content and students' involvement in the classroom (Line 13, I4a, S7b; Line 15, I4a, S7b; Line 18, I4a, S7b). She checked students' understanding individually. Four out of the teachers' seven communicative behaviours were coded as C1a (lines 7, 13, 15 and 17). That means that teachers are very good at respecting students.

Table 7.7 Lesson 2 Learning present tense

Actor	Line	Learning present tense	TA			
			Code1	Code2	Code3	Code4
Teacher	1	Are these examples from the book? Right?	I4a	S7b		
Students	2	Yes.	S7b			
Teacher	3	What are these examples for? Is it for your fun? Right? What are these examples for?	I4a	S7b		
Students	4	(silence).	NA			
Teacher	5	What is the example sentence used for? What is it for? What are the examples in your book used for?	I4a	S7b		
Students	6	(silence).	NA			
Teacher	7	Aren't you full? Can't the brain think? What's the use of examples? It's for demonstration. What's the use of examples?	I4a	S7b		
Students	8	For demonstration.	S7b			
Teacher	9	I've demonstrated it to you. Look at this question. What's the verb form of the question after DID?	I4a	S7b		
Students	10	(silence).	NA			
Teacher	11	Don't you see the example? What's the expression in the example? This is the demonstration. What tense is used? What is the form of a verb?	I4a	S7b		
Students	12	(silence).	NA			
Teacher	13	Come on. You can't tell, can you? What tenses are these words? Use your head.	I4a	S7b		
Students	14	The root form of the verb.	S7b			
Teacher	15	What tenses are these words?	I4a	S7b		
Students	16	The root form of the verb.	S7b			
Teacher	17	So, what words should be used in this place?	I4a	S7b		
Students	18	Go.	S7b			

In LESSON 4, the teacher adopted pictures describing as a learning activity (Table 7.9). Code I3c appeared in every line in this learning activity since the teacher utilised picture materials that connected with students' lives. That raised students' strong interest and initiative in this learning activity. The teacher put forward a series of questions around the given pictures to check the students' understanding of the grammar (Line 128, I4a; Line 130 I4a; Line 132, I4a; Line 134, I4a). Questioning on life connected materials also promote students' participation and allows them to take on a classroom role (S7b). Overall, 13 out of 14 lines were coded with two or three codes. This incident illustrates teacher-student interaction was of high quality in this learning activity.

Teaching styles differ among these four lessons and show a large gap between high-quality and low-quality lessons. The difference between a good lesson and a weak one is noticeable. In outstanding high-quality lessons, teachers respected

Table 7.8 Lesson 3 Reviewing learned vocabularies

Agent	Line	Reviewing learned vocabularies	TA			
			Code1	Code2	Code3	Code4
Teacher	7	All of you, please stand up. OK, first, I will say an English word, and I'll call your name to say the Chinese. OK?	C1a	C2a	I3d	S7b
Students	8	Ok.	S7b			
Teacher	9	If I say home, you should say?	I4a	S7b		
Students	10	回家 (<i>Going home</i>)	S7b			
Teacher	11	Yes. And this time, I will call your name, OK?	C2a	S7b		
Students	12	OK.	S7b			
Teacher	13	If I don't call your name, please be quiet. Let's have a try. So first one, first one. '家' (<i>home</i>), you please.	C1a	C2a	S7b	
Student A	14	Home.	S7b			
Teacher	15	OK, sit down, please. 'Got'.	C1a	C2a	S7b	
Student B	16	得到 (<i>Got it</i>).	S7b			
Teacher	17	OK, sit down, please. Very good.	C1a	C1b	C2a	
Teacher	18	Let's go on. 'Space'.	S7b			
Student C	19	太空 (<i>Space</i>).	S7b			

Table 7.9 Lesson 4 Describing pictures

Agent	Line	Describing pictures	TA			
			Code1	Code2	Code3	Code4
Teacher	126	Who is she?	I3c	S7b		
Students	127	Fu Yuanhui.	I3c	S7b		
Teacher	128	What is she doing?	I3c	I4a	S7b	
Students	129	She is swimming.	I3c	S7b		
Teacher	130	Is she good at swimming?	I3c	I4a	S7b	
Students	131	Yes. She is.	I3c	S7b		
Teacher	132	How old is she?	I3c	I4a	S7b	
Students	133	I don't know.	I3c	S7b		
Teacher	134	Do you like her?	I3c	I4a	S7b	
Students	135	Yes. Maybe.	I3c	S7b		
Student A	136	I like Sun Yang.	I3c	S7b		
Teacher	137	Yeah, you like Sun Yang. Why? Could you?	I3c	I6a	S7b	
Student A	138	Because he is very handsome, and swim well.	I3c	I6c	S7b	
Teacher	139	He can swim very well. But Fu Yuanhui can swim very well too.	I3c			

students, articulated clear expectations, and let students play a role in classroom learning. These are some weaknesses of low-quality lessons. For LESSON 1 and LESSON 2, teachers' behaviours did not show good respect, affecting students'

interest in the lesson. Teachers also did not make their expectations for students on classroom activity clear. This teaching behaviour makes it difficult for students to understand the teacher's intention. In the end, the students could not give the expected responses. Moreover, having no opportunity to play a role in the classroom made students lack participation and fail to learn confidently.

6 Discussion and Conclusion

6.1 Instrument Characteristics as Biases and Limitations

As shown in Table 7.5, only some general teaching behaviours are assessed (I3a to I6c) in TEACH, which means teachers only need to conduct common teaching behaviours to meet the standards to get higher scores.

'High-quality' lessons ranked a little lower in ICALT than in TEACH. It indicated that ICALT has higher overall classroom teaching requirements than TEACH. Regarding 'low-quality lessons ranked higher in ICALT than in TEACH, the teachers in these two classes did not perform well in general teaching behaviour, but they had deeper teaching behaviour. Nevertheless, it does not affect the determination of the final characterisation of 'low-quality.'

Our results indicated that TEACH is a feasible coding scheme for in-depth qualitative analysis on dialogic teaching as it fit our research demands to associate it with a quantitative lesson observation instrument. There is a trade-off between instrument complexity and ease of usage as TEACH was developed to provide quick training for practitioners in developing countries for teacher evaluation and professional teacher development. In contrast, ICALT was initially developed for high-stake inspections and subsequently for high-quality research in developed and developing countries (Maulana et al., 2021).

6.2 The Practicability of Promoting Teacher Reflections: TEACH vs ICALT

The quantitative results indicated that ICALT predicted student engagement better than TEACH. However, the subscale *Learner engagement* is part of ICALT, so it is not surprising that the results might favour ICALT more than TEACH. However, both ICALT and TEACH results showed that clear and structured instructions improve student engagement. Adequate instructions could contribute to a better and depth understanding of classroom activities and contents, resulting in higher student involvement in classroom learning (Boston & Candela, 2018).

Moreover, among the ICALT domains, the average score of the *Adjusting Instructions and Learner Processing to Inter-Learner Differences* was lower than

other domains in ICALT, indicating that teachers in the sample hardly presented student-centred instructions to address learner diversity. A lower rating might be caused by the limited background information of the students available to the raters. The raters did not know the students' learning differences ahead of the class; hence, it might be hard for them to identify students with diverse learning needs to associate teaching behaviours expected to address learner diversity during the classroom observation (Edwards et al., 2006). Thus, a rater may be biased against the teacher if s/he lacks the understanding of students as learners. Among TEACH factors, teachers with better socioemotional skills, including autonomy, perseverance, social and collaborative skills, could have engaged students better in classroom learning.

In addition to the low average score, the *Adjusting Instructions and Learner Processing to Inter-Learner Differences* subscale also has poor reliability. A similar reason that observers lack contextual information in the classroom might affect the reliability. For example, it is not easier to identify whether a student is weaker without asking the teacher. Another explanation is that as the teaching quality of each teacher was assessed based on one single lesson, personalised instruction to fit in inter-learner differences and adjusting might not be readily recognisable in one single lesson but more evident in more lessons observed for the whole academic term. A longitudinal study in which teaching quality can be assessed several times throughout a whole academic term or year could be conducted in the future to better capture student-centred instructions in the teaching quality.

7 Conclusion

Two significant limitations of the present study were the small sample size and selection of samples. In this study, as the sampling only covered teaching whose teaching experience was more than two years and less than eight years, the teachers who taught more than eight years or just started to teach less than two years were underrepresented. Future studies can focus on the assessments and comparisons of teaching quality based on teachers with all lengths of teaching experience. For example, a study on 47 rural primary schools in Guizhou Province showed that the length of teaching experiences varied across teachers, and teachers with 4–10 years of teaching experience only accounted for 27% of the population (Peng, 2015).

Teacher-student interaction is an essential factor affecting classroom teaching quality (Berlin & Cohen, 2018). The differences between high-quality and low-quality lessons are highlighted in respecting students, behaviour expectation for students, and students playing a role in classroom aspects. If a class does not have these characteristics, it is challenging to associate students' interests with specific teaching behaviours and subsequently affect the student learning achievement and make a fair judgement on teaching quality.

There are many classroom observation tools for us to choose for teacher evaluation and research. However, we compared two instruments designed for different purposes and probably for different audiences and contexts. When choosing these

tools, we should first consider comparing the lens of different instruments (Walkington & Marder, 2018; Walkowiak et al., 2019), as we have done to balance efficiency and exhaustivity for the research needs. When analysing the comparative results, we should also thoroughly consider the limitations of our observation tools. We also conducted in-depth qualitative analyses because high-inference classroom observation instruments like ICALT and TEACH cannot provide detailed accounts of classroom processes. Our coding strategies also provide the potential for quantifying qualitative data. We suggest systematic in-depth qualitative analysis with detailed contextual information provide dby the teacher and a longitudinal approach be indispensable to complement high-inference instruments in more objective research and fairer teacher evaluation.

References

- Alexander, R. (2006). *Towards dialogic teaching* (3rd ed.). Dialogos.
- Alexander, R. (2008). *Towards dialogic teaching: Rethinking classroom talk* (4th ed.). Dialogos.
- Anwar, S., & Menekse, M. (2021). A systematic review of observation protocols used in post-secondary STEM classrooms. *Review of Education*, 9(1), 81–120.
- Berlin, R., & Cohen, J. (2018). Understanding instructional quality through a relational lens. *ZDM*, 50(3), 367–379.
- Boston, M. D., & Candela, A. G. (2018). The instructional quality assessment as a tool for reflecting on instructional practice. *ZDM*, 50(3), 427–444.
- Boston, M., Bostic, J., Lesseig, K., & Sherman, M. (2015). A comparison of mathematics classroom observation protocols. *Mathematics Teacher Educator*, 3(2), 154–175.
- Caughlan, S., & Jiang, H. (2014). Observation and teacher quality: Critical analysis of observational instruments in preservice teacher performance assessment. *Journal of Teacher Education*, 65(5), 375–388.
- Chiangkul, W. (2016). *The state of Thailand education 2014/2015 “How to reform Thailand education towards 21st century?”*. Office of the Education Council.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.
- Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Association for Supervision and Curriculum Development.
- Edwards, C. J., Carr, S., & Siegel, W. (2006). Influences of experiences and training on effective teaching practices to meet the needs of diverse learners in schools. *Education*, 126(3), 580–591.
- Erickson, F. (2011). Uses of video in social research: A brief history. *International Journal of Social Research Methodology*, 14(3), 179–189.
- Gersten, R., Baker, S. K., Haager, D., & Graves, A. W. (2005). Exploring the role of teacher quality in predicting Reading outcomes for first-grade English learners: An observational study. *Remedial and Special Education*, 26(4), 197–206.
- Hannum, E. (2005). Market transition, educational disparities, and family strategies in rural China: New evidence on gender stratification and development. *Demography*, 42(2), 275–299.
- Hannum, E., Behrman, J., Wang, M., & Liu, J. (2008). Education in the reform era. In L. Brandt & T. G. Rawski (Eds.), *China’s great economic transformation* (pp. 215–249). Cambridge University Press.
- Hannum, E., Cherng, H. Y. S., & Wang, M. (2015). Ethnic disparities in educational attainment in China: Considering the implications of interethnic families. *Eurasian Geography and Economics*, 56(1), 8–23.

- Heckman, J. J., & Yi, J. (2012). *Human capital, Economic Growth, and Inequality in China* (No. w18100). National Bureau of Economic Research.
- Hennessy, S., Rojas-Drummond, S., Higham, R., Márquez, A. M., Maine, F., Ríos, R. M., García-Carrión, R., Torreblanca, O., & Barrera, M. J. (2016). Developing a coding scheme for analysing classroom dialogue across educational contexts. *Learning, Culture and Social Interaction*, 9, 16–44.
- Hennessy, S., Howe, C., Mercer, N., & Vrikki, M. (2020). Coding classroom dialogue: Methodological considerations for researchers. *Learning, Culture and Social Interaction*, 25, 100404.
- Henry, M. A., Murray, K. S., Hogrebe, M., & Daab, M. (2009). *Quantitative analysis of indicators on the RTOP and ITC observation instruments*. MSPnet. http://mspnet-static.s3.amazonaws.com/MA_Henry_Quantitative_Analysis_RTOP_ITC_112509_FINAL.pdf
- Howe, C., & Abedin, M. (2013). Classroom dialogue: A systematic review across four decades of research. *Cambridge Journal of Education*, 43(3), 325–356.
- Howe, C., & Mercer, N. (2017). Commentary on the papers. *Language and Education*, 31(1), 83–92.
- Howe, C., Hennessy, S., Mercer, N., Vrikki, M., & Wheatley, L. (2019). Teacher-student dialogue during classroom teaching: Does it really impact on student outcomes. *Journal of the Learning Sciences*, 28(4–5), 462–512. <https://doi.org/10.1080/10508406.2019.1573730>
- Huijgen, T., van de Grift, W., Van Boxtel, C., & Holthuis, P. (2017). Teaching historical contextualisation: The construction of a reliable observation instrument. *European Journal of Psychology of Education*, 32(2), 159–181.
- Jacobs, J. K., Kawanaka, T., & Stigler, J. W. (1999). Integrating qualitative and quantitative approaches to the analysis of video data on classroom teaching. *International Journal of Educational Research*, 31, 717–724.
- Klette, K. (2009). Challenges in strategies for complexity reduction in video studies. Experiences from the PISA+ study: A video study of teaching and learning in Norway. In T. In Janík & T. Seidel (Eds.), *The power of video Studies in investigating teaching and learning in the classroom* (pp. 61–82). Waxmann.
- Ko, J. Y. O. (2010). *Consistency and variation in classroom practice: A mixed-method investigation based on case studies of four EFL teachers of a disadvantaged secondary school in Hong Kong* (Publication No. 11363). Doctoral dissertation, University of Nottingham. Nottingham eThesis.
- Ko, J., Ho, M., & Chen, W. (2015, July–Oct). *Teacher report —teacher effectiveness and goal orientation. report Nos. 1–43*. Tai Po: Hong Kong Institute of Education.
- Kyriakides, L., Creemers, B. P., & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and teacher education*, 25(1), 12–23.
- Lesh, R. A., & Lehrer, R. (2000). Iterative refinement cycles of videotape analyses of conceptual change. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 665–708). LEA.
- Li, W., Park, A., Wang, S., & Jin, L. (2007). School equity in rural China. In E. Hannum & A. Park (Eds.), *Education and reform in China* (pp. 27–43). Routledge.
- Marshall, J. H., & Sorto, M. A. (2012). The effects of teacher mathematics knowledge and pedagogy on student achievement in rural Guatemala. *International Review of Education*, 58(2), 173–197.
- Marshall, J. C., Smart, J., Lotter, C., & Sirbu, C. (2011). Comparative analysis of two inquiry observational protocols: Striving to better understand the quality of teacher-facilitated inquiry-based instruction. *School Science and Mathematics*, 111(6), 306–315.
- Maulana, R., André, S., Helms-Lorenz, M., Ko, J., Chun, S., Shahzad, A., et al. (2021). Observed teaching behaviour in secondary education across six countries: Measurement invariance and indication of cross-national variations. *School Effectiveness and School Improvement*, 32(1), 64–95.
- McDonald, R. P. (2013). *Test theory: A unified treatment*. Psychology Press.

- Mercer, N. (2010). The analysis of classroom talk: Methods and methodologies. *British Journal of Educational Psychology*, 80(1), 1–14.
- Molina, E., Fatima, S. F., Ho, A. D. Y., Melo Hurtado, C. E., Wilichowski, T., & Pusharatnam, A. (2018). *Measuring teaching practices at scale: Results from the development and validation of the teach classroom observation tool* (World Bank Policy Research Working Paper, No. 8653). The World Bank.
- Muijs, D., & Reynolds, D. (2005). *Effective teaching: Introduction & conclusion*. Sage.
- Muijs, D., Kyriakides, L., Van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art—teacher effectiveness and professional learning. *School Effectiveness and School Improvement*, 25(2), 231–256.
- OECD. (2019, April 26). *PISA 2018 assessment and analytical framework*. https://www.oecd-ilibrary.org/education/pisa-2018-assessment-and-analytical-framework_b25efab8-en
- Peng, Y. (2015). *The recruitment and retention of teachers in rural areas of Guizhou, China*. Doctoral dissertation, University of York, York, UK.
- Pianta, R. C., & Hamre, B. K. (2009). Conceptualisation, measurement, and improvement of classroom processes: Standardised observation can leverage capacity. *Educational Researcher*, 38(2), 109–119.
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system™: Manual K-3*. Paul H Brookes Publishing.
- Qian, X., & Smyth, R. (2008). Measuring regional inequality of education in China: Widening coast–inland gap or widening rural–urban gap? *Journal of International Development: The Journal of the Development Studies Association*, 20(2), 132–144.
- Sammons, P. (2007). *School effectiveness and equity: Making connections*. CfBT.
- Sapire, I., & Sorto, M. A. (2012). Analysing teaching quality in Botswana and South Africa. *Prospects*, 42(4), 433–451.
- Schoenfeld, A. H. (2013). Classroom observations in theory and practice. *ZDM*, 45(4), 607–621.
- Shi, Y., Shen, X., Wang, T., Cheng, L., & Wang, A. (2021). Dialogic teaching of controversial public issues in a Chinese middle school. *Learning, Culture and Social Interaction*, 30, 100533.
- Song, Y., Chen, X., Hao, T., Liu, Z., & Lan, Z. (2019). Exploring two decades of research on classroom dialogue by using bibliometric analysis. *Computers & Education*, 137, 12–31.
- Stearns, L. M., Morgan, J., Capraro, M. M., & Capraro, R. M. (2012). A teacher observation instrument for PBL classroom instruction. *Journal of STEM Education: Innovations & Research*, 13(3), 7.
- Stigler, J. W., Gonzales, P., Kwanaka, T., Knoll, S., & Serrano, A. (1999). The TIMSS videotape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States. National Center for Education Statistics (ED), .
- Teddle, C., Creemers, B., Kyriakides, L., Muijs, D., & Yu, F. (2006). The international system for teacher observation and feedback: Evolution of an international study of teacher effectiveness constructs. *Educational Research and Evaluation*, 12(6), 561–582.
- Tee, M. Y., Samuel, M., Nor, N. M., Sathasivam, R. V., & Zulnaidi, H. (2018). Classroom practice and the quality of teaching: Where a nation is going? *Journal of International and Comparative Education*, 7(1), 17–33.
- Tsang, M. C., & Ding, Y. (2005). Resource utilisation and disparities in compulsory education in China. *China Review*, 5, 1–31.
- Van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research*, 49(2), 127–152.
- Vrikki, M., Wheatley, L., Howe, C., Hennessy, S., & Mercer, N. (2019). Dialogic practices in primary school classrooms. *Language and Education*, 33(1), 85–100. <https://doi.org/10.1080/09500782.2018.1509988>
- Walkington, C., & Marder, M. (2018). Using the UTeach Observation Protocol (UTOP) to understand the quality of mathematics instruction. *ZDM*, 50(3), 507–519.

- Walkowiak, T. A., Adams, E. L., & Berry, R. Q. (2019). Validity arguments for instruments that measure mathematics teaching practices: Comparing the M-Scan and IPL-M. In J. D. Bostic, E. E. Krupa, & J. C. Shih (Eds.), *Assessment in mathematics education contexts* (pp. 90–119). Routledge.
- Wang, J., & Li, Y. (2009). Research on the teaching quality of compulsory education in China's west rural schools. *Frontiers of Education in China*, 4(1), 66–93.
- World Bank. (2019, July 9). *TEACH manual*. World Bank. https://www.worldbank.org/en/topic/education/brief/teach-helping-countries-track-and-improve-teaching-quality?cid=EXT_WBEmailShare_EXT
- Wragg, T. (2013). *An introduction to classroom observation*. Routledge.
- Yang, J., Huang, X., & Li, X. (2009). Education inequality and income inequality: An empirical study on China. *Frontier of Education in China*, 4(3), 413–434.
- Yang, J., Huang, X., & Liu, X. (2014). An analysis of education inequality in China. *International Journal of Educational Development*, 37, 2–10.
- Zeng, J., Pang, X., Zhang, L., Medina, A., & Rozelle, S. (2014). Gender inequality in education in China: A meta-regression analysis. *Contemporary Economic Policy*, 32(2), 474–491.
- Zhang, H. (2017). Opportunity or new poverty trap: Rural-urban education disparity and internal migration in China. *China Economic Review*, 44, 112–124.

Jieyan Celia Lei is a lecturer at Shaoyang University in Mainland China and a doctoral candidate at the Education University of Hong Kong, where she graduated with an MEd degree and worked as a research assistant at the Department of Educational Policy and Leadership. She published on teaching quality of public schools in an inland province in Mainland China and has extensive research experience using various classroom observation instruments. Her research interests include dialogic teaching, teaching quality, and metacognitive teaching.

Zhijun Chen is a Doctoral Researcher at the Department of Education at the University of Bath (UK). She holds an MSc in Psychology from the University of St Andrews (UK). She also works as a research assistant with Dr James Ko at the Education University of Hong Kong, focusing on teaching quality and teaching assessment in multiple cultures. Her research interests include educational effectiveness, large-scale international assessments (PISA, TIMSS, PIRLS, ERCE, etc.), education inequality, and classroom observation.

Dr James Ko is an Associate Professor at the Department of Policy Leadership and Co-Director of the Joseph Lau Luen Hung Charitable Trust Asia Pacific Centre for Leadership and Change at the Education University of Hong Kong. Before his doctoral study, James was an EFL teacher for about 20 years and led two functional teams in a secondary school for 10 years. He is a recurrent grantee of the RGC and UGC grants and the principal investigator of 23 projects, collaborating with local academics and overseas researchers on 40 projects. He has supervised 14 doctoral students with 8 completed.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 8

Measuring Teaching Skill of South Korean Teachers in Secondary Education: Detecting a Teacher's Potential Zone of Proximal Development Using the Rasch Model



Wim van de Grift, Okhwa Lee, and Seyeoung Chun

Abstract Many observation instruments are in use to make the skills of teachers visible. These tools are used for assessment, for guidance and coaching, and for policy-oriented research into the quality of education. Depending on the purpose of use of an observation instrument, we not only need more observations about the same teacher, but the observation instrument must also meet higher psychometric requirements. Observation instruments only used to assess sample characteristics, such as the mean and dispersion, require less stringent psychometric requirements than observation instruments that are used to assess individuals. For assessing sample characteristics, it is also not necessary to do more than one observation with each respondent. Observation instruments used for individual assessments that lead to high stake decisions should meet the highest psychometric requirements possible. We can slightly mitigate the psychometric norms attached to an observation tool that is only used for guidance and coaching on the condition that the observed teacher explicitly informed that the observed lesson was representative and that this lesson offered sufficient opportunities to demonstrate all the skills the teacher has. Nevertheless, there are also additional requirements that must be met by observation instruments that are used for guidance and coaching. For good guidance and coaching, it is usually not very useful to tell an observed teacher only what went right or wrong. Teachers need concrete instructions to be able to improve. Many things that have not gone very well are often (and sometimes far) out of the reach of the teacher being observed. Coaching skills that are beyond the reach of the observed person will lead to disappointment rather than to the desired effect. The important thing in

W. van de Grift (✉)

Emeritus Professor of Education, University of Groningen, Groningen, The Netherlands

O. Lee

Department of Education, Chungbuk National University,
Cheongju, South Korea

S. Chun

Department of Education, Chungnam National University, Daejeon, South Korea

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_8

good guidance and coaching is to ensure that the observed teacher is going to take that very step, that is within his reach, but that he has not just set. Then, of course continue with the next steps, leading to incremental progress. For this, we need to have an insight into the successive difficulty of the different skills of teachers. In the past, we gained some experience with the use of the Rasch model to gain an insight into the successive level of difficulty in the actions of Dutch teachers working in elementary education. These studies are all done with the **I**nternational **C**omparative **A**nalysis of **L**earning and **T**eaching (ICALT) observation instrument. In this chapter, we are trying to make a next step by using the Rasch model for detecting the zone of proximal development of the observed teachers. Another new element in this study is the following: Until now, the ICALT observation instrument has been used mainly in (the culture of) European schools. In this chapter, we focus on Asian secondary education, as it takes shape in South Korea.

Keywords Teaching skill · Zone of proximal development · Rasch model

1 Introduction

Many observation instruments are in use to make the skills of teachers visible (cf. Bell et al., 2018; Dobbelaer, 2019). These tools are used for assessment, for policy-oriented research into the quality of education and for guidance and coaching. For good guidance and coaching, it is usually not very useful to tell an observed teacher only what went right or wrong. Teachers need concrete instructions to be able to improve. Many things that have not gone very well are often (and sometimes far) out of the reach of the teacher being observed. Coaching skills that are beyond the reach of the observed person will lead to disappointment rather than to the desired effect. The important thing in good guidance and coaching is to ensure that the observed teacher is going to take the next step, within his or her reach, that s/he has not yet reached. After that the following steps can be taken, leading to incremental growth. For this, we need to have an insight into the successive difficulty of the different skills of teachers. In this article, we use the Rasch model (Rasch, 1960, 1961) for detecting the potential zone of proximal development of the observed teachers.

The observation instrument we will use is the ICALT instrument. The ICALT observation instrument was developed between 1989 and 1994 for primary education and was initially used by the Education Inspectorate (Van de Grift & Lam, 1998). The instrument, which has also been used by other European education inspectorates (cf. Van de Grift, 2007, 2014), currently has a version consisting of six Likert scales. The six Likert scales contain 32 high inferential items and 120 low inferential examples of good practice. The 152 high and low inferential items are all based on reviews of a large number of studies on the effectiveness of education on student achievement (cf. the references). The 32 high inferential items are the core of the observation instrument. The raw score on the instrument is simply the sum score on these 32 items. These 32 items have an abstract or high inferential

character. An example of a high inferential item is "... promotes learners' self-confidence". In the observation instrument, every high inference item is accompanied by several low inference items. For example, low inferential items that belong to the high inferential item above are "...gives positive feedback on questions and remarks from learners", "...compliments learners on their work", and "...acknowledges the contributions that learners make". The actions in these low inferential items are coded as simply observed or not observed during a lesson. The 120 low inferential items are used in different situations. During the training of the observers, the low inferential items are used to explain the height of the score on the 32 high inferential items. If the score on a high inferential item is low, the scores on the corresponding low inferential items should also be low. When an observer gives a low score on a high inferential item, the scores on the corresponding low inferential items should also be low. Also, the scores of the low inferential items are used when coaching the observed teacher. It has little practical value to use the abstract and high inferential items for that. It is more informative for the observed teacher when the advice based on the low inference items is: 'evaluate whether the lesson aims have been reached' and 'offer weaker learners extra study and instruction time', than the advice based on the high inference item 'adjust instructions and learner processing to inter-learner differences'. The low inference items indicate more concretely what the observed teacher should do. (For more details, see the appendix with the ICALT instrument.)

The first three Likert scales concern the basic skills of teaching: creating a safe and stimulating educational learning climate, organizing the lesson efficiently, and providing clear and structured instruction. The other three Likert scales concern the advanced teaching skills: giving an intensive and activating lesson, tailoring instruction and processing to differences between students and teaching students learning strategies. An observed teacher masters the observed activities from a scale to a more than sufficient extent when the score in that domain is higher than 2.5. (Then $\geq 65\%$ of the items is scored sufficient.) The six domains of the ICALT instrument show a hierarchical order with increasing difficulty (Van de Grift, 2021). The items from some domains of the observation instrument are relatively easy for teachers to master, for example creating a safe and stimulating learning environment. Other domains are relatively difficult for teachers, for example differentiated teaching and teaching students learning strategies. This hierarchical order in the domains of the ICALT instrument made us wonder whether this order could also be found in the individual items. Therefore we studied in a sample of 400 teachers working with 6–12-year-old students the question whether the 32 individual items meet the requirements of the dichotomous Rasch model. We found a reliable Rasch scale with 31 items for measuring the teaching skills. The simplest items concerned basic skills such as creating a safe learning environment, efficient classroom management and clear and structured instruction. The slightly more difficult items concerned activating learners. The items concerning differentiated instruction were clearly more difficult. The most difficult items were those related to teaching students how to learn. The scale is suitable for distinguishing six zones that give an indication of the zone of proximal development of an observed teacher (Van de Grift et al., 2019).

In 2008, we began studies to determine whether the ICALT observation instrument could also be used reliably and validly with student teachers and beginning teachers in secondary education (Maulana et al., 2015, 2016). In 2015, we started international comparisons of the quality of teaching in various non-Western countries, such as South Korea (Van de Grift et al., 2017) and South Africa (De Jager et al., 2017). In the same period we started analyses in which we investigated whether the Rasch model was applicable to the pedagogical didactic behaviour of teachers in secondary education (Van de Grift et al., 2014; Van der Lans et al., 2017, 2018). The order of the difficulty of the 31 items that fitted the Rasch model appeared to be more or less the same for teachers in secondary education as it was for teachers in basic education. The simplest items concerned basic skills such as creating a safe learning environment, efficient classroom management and clear and structured explanations. The slightly more difficult items concerned activating students. Clearly more difficult were the items about teaching pupils how to learn. In contrast to the situation in primary education, the items that concerned the provision of differentiated instruction proved to be the most difficult in secondary education. The fact that the items providing differentiated instruction were the most difficult for teachers in secondary education probably has to do with the fact that students in primary education are not sorted by skills level as they are in secondary education. In the present publication, we investigate whether this order item difficulties is maintained among secondary school teachers from a completely different culture, the Asian culture.

2 Theoretical and Empirical Background

In this section, we will introduce the idea of “zone of proximal development”.

After that we will go into some theoretical and empirical backgrounds of

- the relationships between teaching skills and students’ learning gain
- the trainability of teaching skills, and
- the relationships between the growth of teaching skills and growth in students’ learning gain.

2.1 *The Idea of the “Zone of Proximal Development”*

Many years ago, the concept “zone of proximal development” was introduced by Vygotsky (1930). Vygotsky was interested in the ontogenetic (and phylogenetic) development of thinking and speech. In his conception the zone of proximal development relates to the difference between what a child can achieve independently

(the so-called actual level of development) and what a child can achieve with guidance and encouragement from a skilled person (the so-called zone of proximal development). Over the years, there has been a lot of discussion about the interpretation of the work of Vygotsky. Part of this discussion has to do with the correct translation of several concepts from Russian into western languages (Lompscher & Rückriem, 2002).

Without going in too much detail, we will interpret in this study this concept as an area of learning that is very near to the actual level of skill of a person. We suppose that students, taught in their zone of proximal development, will learn faster and more effectively, than students who are asked to do things that are (too) difficult for them. For example in the teaching of pupils we do not start with an explanation of multiplication before the idea of repeated addition is well understood. We do not start reading comprehension before the child can perform the technical reading process. The zone of proximal development helps to properly determine the upper limit of what a person is already capable of. This is the starting point for feedback and deliberate training and behavioural practice with the aim to raise the upper level of performance to a (slightly) higher level of the proximal development.

In this study, we are interested in the professional development of teachers. The professional development of teachers differs from ontogenetic theories, but there are related matters. An important related matter is the fact that mastering basic knowledge and skills of teaching is conditional for the mastering of more complex knowledge and skills. Research showed that teaching skills associated with differentiation in teaching are more difficult than those related to activating students are. Activating students is more difficult compared to classroom management skills (Van de Grift et al., 2014, 2019; Maulana et al., 2016). Mastering of the basic skills of teaching seems to be conditional for being able to master other more complex teaching skills. Teachers still having problems with classroom management should not be coached in skills to activate students. They should first be helped with their classroom management problems. The same is for teachers who have problems with giving clear explanations; they are not yet ready for differentiated instruction. They must first learn to explain clearly and in a structured way before they can help pupils with specific learning needs.

The one who is in charge of the guidance or coaching of teachers should consider not only the actual level of development but also the zone of proximal development of teachers. The difference between the teachers actual level of development and the level of performance that he or she achieves in collaboration with the coach, defines the zone of proximal development. Coaching of teachers is maximally productive only when it occurs at a certain point in the zone of proximal development. The zone of proximal development determines the domain of improvements that are accessible to the teacher.

However, determining the zone of proximal development of teachers' teaching skills is not a simple and easy task. It is therefore not surprising that the knowledge about this in the current literature is very scarce.

2.2 *Teaching Skills and Students' Learning Gains*

Between 1983 and 2008 several reviews are published about the relationships between teaching behaviour and student achievement. These research reviews make clear that several teaching behaviours are indeed related to student achievement and learning gains: Setting targets, offering sufficient learning and instruction time, monitoring students' achievements, creating special measures for struggling students, establishing a safe and stimulating educational climate, organizing efficient classroom management, giving clear and structured instruction, organizing intensive and activating teaching, differentiating instruction, and teaching learning strategies. Good readable summaries of various reviews of these studies can be found in Marzano (2003) and Hattie (2009, 2012). More detailed information can be found in the references of this chapter. Several econometric studies indicated also that better teachers have students with more learning gains (Hanushek & Rivkin, 2010; Kane & Staiger, 2008; Rivkin et al., 2005).

Some of these teaching behaviours are susceptible to observation; other behaviours have to be found through interviews. In this study, we concentrate on the issues that can be observed by external observers in classes: establishing a safe and stimulating educational climate, organizing efficient classroom management, giving clear and structured instruction, organizing intensive and activating teaching, adapting instruction, and teaching learning strategies.

An important question is: How malleable and trainable is this behavior? The following paragraph deals with this.

2.3 *Trainability of Teaching Skills*

Kraft et al. (2018) reviewed 60 American, Canadian, and Chilean empirical studies on the effects of the coaching of teachers and conducted meta-analyses to estimate the mean effect of coaching programs on teachers' instructional practice. There are 55 American, and 5 Canadian and Chilean empirical studies. The mean effect across 60 studies, employing causal research designs was a pooled effect size of 49% of a standard deviation on teachers' instructional practice.

Van den Hurk et al. (2016) studied 110 teachers, working in Dutch elementary education. These teachers had been coached based on a lesson observed with them. After the coaching these teachers showed a skill growth, on several observed aspects of teaching. They found for creating a safe and stimulating climate a growth of 29% of a standard deviation; for efficient classroom management a growth of 37%; for clear and structured instruction a growth of 62%; for activating students 76%; for teaching learning strategies 71%, and for differentiation they found a growth of 51% of a standard deviation. These Dutch results are in agreement with the average effect size found in the American, Canadian and Chilean studies found by Kraft et al. (2018).

The following section handles the relationship between growth in teaching skills and (extra) growth in student achievements.

2.4 Growth of Teaching Skills and Students' Learning Gains

Kraft et al. (2018) found a mean effect of growth in teaching on student achievement of 18% of a standard deviation. Effect sizes were larger (34% of a standard deviation) in smaller programs than in larger programs (10% of a standard deviation). Therefore, it seems that an average growth of 49% of a standard deviation on teachers' instructional practice in USA, Canada and Chile goes along with an average growth of 18% in students' academic achievement.

In several small-scale experiments done in Dutch elementary education (Houtveen & Van de Grift, 2007a, b; Houtveen et al., 2004, 2014) an average effect size of 64% of a standard deviation was found in the growth of teaching skills by specially observed and coached teachers. The students in the experimental groups of these experiments had an extra learning gain of 45% of a standard deviation for decoding, 38% for comprehensive reading and 52% for mathematics. Therefore in these studies, a growth of almost two third of a standard deviation in teaching skill goes along with a growth of student achievement of almost half a standard deviation.

3 Aim of This Study

We have already seen that 31 of the 32 items of the ICALT observation instrument have a hierarchical order. This hierarchical order is very important for accurately tracing the zone of close development of an observed teacher. In this study, we investigate whether the order of item difficulty found among Dutch secondary school teachers is maintained among secondary school teachers from a totally different culture, the South Korean culture.

4 Method

4.1 Sample Characteristics

In South Korea, the teaching skills of a sample of 375 teachers working in 26 secondary schools in the regions Deajeon, Chungnam, Cheongju, and Chungbuk were observed in one real life lesson by specially trained observers. Teachers in the sample were recruited by their voluntary participation in the research project. They were introduced about ICALT and invited by the observers who had been trained with ICALT tool. These data were previously used in Van de Grift et al. (2017). These 375 teachers taught 25 different subjects. The teachers had, on average, 11 years of teaching experience. About 51% of the teachers were female. The average class size was 29 students (see Table 8.1 for more detailed information).

Table 8.1 Sample characteristics (n = 375 teachers)

Subject	% teachers		Years of experience	Class size
Language	17.9	Mean	11.32	29.12
English	20.5	Standard dev.	9.59	7.17
Beta (math, science, information science and so on)	34.7	Minimum	0	10
Else	26.9	Maximum	38	42

This sample of 375 teachers is large enough to estimate proportions in the population of the regions Deajeon, Chungnam, Cheongju, and Chungbuk with a precision of 5% and a confidence interval of 95% (cf. Kirby et al., 2002). These teachers were observed by 40 trained observers; 14 observers observed <5 lessons and 26 observers observed 9–33 lessons. The observers had on average almost 26 years of experience as a teacher.

4.2 *Translation of the Observation Instrument and Training of Observers*

4.2.1 Translation of the Observation Instrument

The English version instrument was firstly translated into Korean by one of the Korean authors of this chapter. This first translation was back-translated into English from Korean by a native English teacher who were teaching English at a secondary school in South Korea. The back-translated English instrument was examined by both the Dutch ICALT research team and the original Korean translator. Then the Korean version of the instrument had been finalized.

4.2.2 Training of Observers

The observers who participated in this study were trained over the course of two full days. The training involved explanations of the theoretical, empirical and practical backgrounds of the observation instrument, practices with observing two videotaped lessons, and a discussion about how to evaluate teaching behaviours using the associated scoring procedures. Both videotaped lessons were in English.

During the presentation of both video tapes, the observers had to score both high and low inferential items.

After presenting the consensus results of the first video to the observers, discussions were organized between observers who did not agree on one more items. The scores on the low inferential items were used to reach consensus on the scoring of the high inferential items. The scores on the low inferential items are the ‘arguments’ for the score on the high inferential items. These arguments are used during

the discussions. Furthermore, the consensus within the observers and the expert norm was compared, with a cut-off of 0.70. In the current group, the consensus level was 0.82. Only certified observers were invited to observe classrooms.

4.3 Interrater Reliability

It sounds quite simple and reasonable: observers observing the same lesson should reach, working with the same observation instrument, the same conclusion. In order to reach this goal observers should be very consistent with each other in their judgements. Consistency alone is not enough. Observers must also have a high degree of agreement in their scores. Their amount of consensus must also be higher than can be achieved only by guessing.

Several statistics are used to determine whether observers interpret the same event in the same way. Ten Hove et al. (2018) showed that working with the same data, different coefficients show different results. These partially overlapping statistics all have their own merits and advantages, and problems and disadvantages. That is why we use several statistics in this study to obtain an indication of interrater reliability. The results we found with three of these statistics are presented in Table 8.2.

4.3.1 Intra-Class Correlation

We used the intra-class correlation coefficient (ICC; Hallgren, 2012) in order to assess the degree that observers showed consistency in their ratings of teaching skill across the items of the ICALT-scale. According to Cicchetti (1994) the interrater reliability is poor for ICC values less than .40, fair for values between .40 and .59, good for values between .60 and .74, and excellent for values between .75 and 1.0. During the observation training, we used the two video lessons: an English lesson and a geography lesson.

For the English lesson, an ICC of .90 was found, indicating that the observers had a high degree of consistency in their judgements. Studying changes in the ICC when one or more observers were deleted resulted in the conclusion that not inviting two observers should lead to ICC's of respectively .902 and .904. These improvements are not visible when rounded to the second decimal place. Therefore, we had no reason not to invite these observers to continue with this study.

Table 8.2 Coefficients for interrater reliability

	Video English lesson	Video Geography lesson
Intra-class correlation	.90	.95
Percentage agreement	75.14	82.22
Fleiss' κ	.27	.46

For the geography lesson, an ICC of .95 was found, again indicating that the observers had a high degree of consistency in their judgements. In comparison with the first lesson (the English lesson), this is not a major improvement. Looking at the intra-class correlation coefficient, the observers appeared to agree with each other very consistently.

Consistency in ratings is the tendency for one observer to increase, or decrease as another observer increases or decreases. The covariance between the observers plays a very important role in this statistic. This has the disadvantage that strict observers can have high correlations with more indulgent observers, while strict observers nevertheless give more insufficient scores than more lenient observers. That is why we also computed the percentage of agreement between the observers.

4.3.2 Agreement Percentage

A simple and popular method for calculating inter-assessor reliability consists in calculating the percentage agreement of the observers. This is done by adding up the number of items that received identical ratings by the observers and dividing that number by the total number of items rated by observers (Stemler, 2004). The consensus percentage among observers was 75.1% for the English lesson and 82.2% for the geography lesson. This means that the exact agreement on the question sufficient or insufficient was on average over 75% and 82%. This result indicates that the average agreement percentage of the observers is satisfactory.

The highest agreement percentages are found for both the most difficult and most easy items. The relatively low agreement percentages are found around the sum score of the scale. As we will see in paragraph 5.4, the items with the lowest percentages of consensus are exactly in the area of current development of the observed teacher. It is hardly surprising that the exact marking of the skill of the observed teacher causes relatively most consensus problems between the observers.

Several researchers are of the opinion that the percentage of agreement should be corrected for the chance of accidental agreement (Cohen, 1960; Kundel & Polansky, 2003; Landis & Koch, 1977). This is the subject of the following section.

4.3.3 Fleiss' κ

Fleiss' κ is a measure of the agreement between more than two observers, where agreement due to chance is factored out (Cohen, 1960; Fleiss & Cohen, 1973; Fleiss, 1981). Fleiss' κ varies from -1 (perfect disagreement), 0 (no different to change) to 1 (perfect agreement). According to Landis and Koch (1977) the interrater reliability is poor for values less than .00, slight for values between .0 and .20, fair for values between .21 and .40, moderate for values between .41 and .60, substantial for values between .61 and .80, and almost perfect for values between .81 and 1.0 . These intervals for Fleiss' κ are cited as norms in many articles (e.g. Viera & Garret, 2015). Landis and Koch (1977), however, are much more modest in their

article. They are looking for a “consistent nomenclature”. They call their intervals arbitrary. The intervals can be seen as “benchmarks” for the discussion about one of their tables in their article (Landis & Koch, 1977, 165). In their article, Landis and Koch do not provide any empirical arguments for their intervals and their indications of the strength of the agreement.

Falotico and Quatto (2015) found that Fleiss’ κ statistic behaves inconsistently in cases of strong agreement between observers, since this statistic assumes lower values than it would have been expected. In the formula for Fleiss’ κ all items are assessed equivalent. However, in a Rasch scale, the items are not equivalent. Some items are at the beginning of the dimension and are dominated by many teachers. The consensus between observers will be high in that part of the scale. The same applies to the items at the end of the dimension of a scale. Here too the consensus will be high, because many teachers do not meet these items. However, exactly at the point where the current skill of the observed teacher lies, the consensus will be relatively low. If it is important to control for chance, then there must also be a control for the skill level of an observed teacher, otherwise the Fleiss will underestimate.

It would be useful if an empirical study were to be conducted, in which the ‘standards’ of Landis and Koch would be validated. This is also done by Lipsey (1990) for the standards that Cohen (1967) proposed for effect size differences.

We started the observation training with video about an English lesson. On this video, we found a Fleiss’ κ of .27, indicating a fair agreement (according to Landis and Koch) between the observers. For the geography lesson, we found a Fleiss’ κ of .46, indicating a moderate agreement (according to Landis and Koch) between the observers. In view of the discussion above, we are inclined that the Fleiss’ kappa’s, we found make it clear in any case, that the agreement found between the observers is not based on chance only.

We found that after the training the observers grew in their mutual consistency and their degree of agreement. The extent to which their agreement could be explained by chance alone decreased after the training.

Furthermore, we found that observers were very consistent with each other in their judgments. The observers also had a high degree of agreement in their scores. Their amount of consensus was higher than can be achieved by guessing alone.

Each of the observers was invited to participate in this study. We may conclude that these results are sufficient to set up a study into the characteristics of the frequency distribution in the sample.

For a study in which we want to determine the area of immediate development of individual teachers, the ICC is sufficiently high, but it is also important that the percentage of agreement of the items in the middle of the Rasch scale is at least 70%.

4.4 The Fit of the Rasch Model

In a Guttman (1950) scale, items are arranged in such an order that an individual who responds correctly on a particular item also respond correctly on items of lower rank-order. With the perfect Guttman scale one is able to predict with the raw score

alone, which items were responded correctly or not. To measure a person's ability, Guttman scale is very helpful for finding a person's zone of proximal development. This "deterministic" Guttman model, however, works fine for constructs that are strictly hierarchical and highly structured. In most social science contexts however, data from respondents often do not closely match Guttman's deterministic model. That is why Guttman's deterministic model is brought within the probabilistic framework of the Rasch model. The Rasch model (Rasch, 1960, 1961) offers unique possibilities for arranging items and persons on a single dimension. Item difficulty parameters and abilities of persons can be estimated independently and find their location on the same dimension. The Rasch model requires the data of a scale to satisfy three assumptions:

- the scale should be unidimensional,
- the items of the scale should be local stochastic independent, and
- the item characteristic curves should be parallel.

We therefore checked whether the evaluations of the observers made with this instrument met these assumptions.

In most cases, a measurement scale is only used to determine the score of a person, because we are interested in the sample mean. In our case however, we are less interested in the average score of a sample. In our study, we are concerned with the scores of individual teachers in order to be able to coach them. This means that we have to set higher requirements in the quality of the individual items. That means also that we cannot work with global testing alone. We also need to map the quality of individual items. This requires tests that provide a detailed picture of the functioning of the individual items. Therefore, model-data fit analyses will be carried out using several different statistical programs.

Another reason for using different analysis techniques is that many analysis techniques do not really provide **the** proof, or **the** hard evidence for unidimensionality, local independence or parallelism of item characteristic curves.

4.4.1 Unidimensionality

The assumption of unidimensionality states that observations can be ascribed to a single latent construct, in our case: teaching skill observable in the classroom. The unidimensionality assumption of a (Rasch) scale is difficult to confirm or to disconfirm (DeMars, 2010). Nevertheless, we can use several procedures to test whether it is likely that a set of items form a unidimensional scale.

Confirmatory Factor Analysis

A possible procedure is using confirmatory factor analysis (CFA) with a one-factor model. For this analysis, we used the program Mplus 7.4 (Muthen & Muthen, 1998–2015). The usual χ^2 -based test for model fit is substantially affected by

sample size (Marsh et al., 1988). Because we have a large sample of observations, we use the Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI). Both indices are less vulnerable to sample size. Furthermore, we consider the Root Mean Square Error of Approximation (RMSEA) to assess model fit. The norms for acceptable fit are CFI and TLI $> .90$ and RMSEA $< .08$ (Chen et al., 2008; Hu & Bentler, 1999; Marsh et al., 2004; Kline, 2005; Tucker & Lewis, 1973; Cheung & Rensvold, 2002).

Table 8.3 shows that both the Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI) for the dichotomised 32 items are above the norm of .90 and the Root Mean Square Error of Approximation (RMSEA) is below the norm of .08, which is an indication for unidimensionality.

In order to determine whether the one-factor model is an optimal model, we investigated whether a four-factor model that corresponds to the areas of proximal development found (cf. Table 8.9) might be a better alternative. This was not the case. Both the CFI and the TLI of this four-factor model were unacceptably low (respectively .728 and .708) and the RMSEA of this four-factor model was .132, which is unacceptably high (cf. Table 8.3).

A Scree Plot of Eigenvalues

Another way to check whether the 32 items of the teaching skill together form a unidimensional latent construct is using a “graphical test” by making a scree plot of the eigenvalues based on the correlation matrix of items. The eigenvalues of the factor analysis are plotted in Fig. 8.1.

The first eigenvalue (11.23) is considerably larger than the second (1.86) and third (1.49) eigenvalues. These results indicate that the scree plot clearly shows one dominant factor, which indicates that the assumption of unidimensionality seems to be reasonable.

Factor analysis is an analysis technique that stems from the classical test theory. Factor analysis is based on the factor loadings of the items. In the Rasch model, not so much the factor loadings as the item difficulties play a central role. That is why we need to extend the research into unidimensionality of the Rasch scale with a technique that has been specially developed for the Rasch model. We will use Andersen’s (1973, 1977) log-likelihood ratio test. This analysis technique developed by Andersen also offers excellent possibilities to trace the items that cause disruptions of the unidimensionality.

Table 8.3 Confirmatory factor analyses

	CFI	TLI	RMSEA
Norms for model fit	$>.90$	$>.90$	$<.08$
Results of the intended one-factor model	.964	.961	.048
Results of an alternative four-factor model	.728	.708	.132

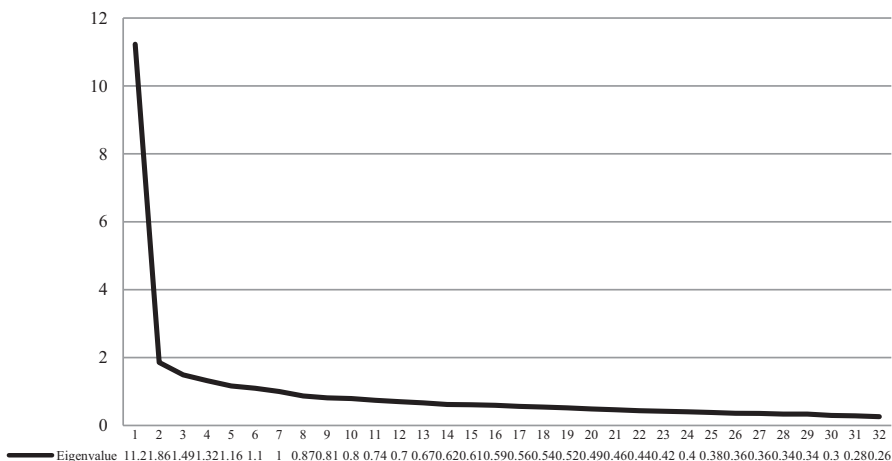


Fig. 8.1 Scree plot of eigenvalues

Table 8.4 Anderson’s log likelihood ratio test for different teacher characteristics

	Anderson’s χ^2	df	p-value
Gender	45.781	31	.042
Gender leaving out item 27	39.282	30	.120
Teaching experience (<5 years of experience and \geq 5 years of experience)	35.812	31	.253
β -Subject matter (math, science, information science and so on) versus language, English and other subject matters	37.526	31	.195
Class size (<30 students and \geq 30 students)	41.174	31	.105

Anderson’s Log Likelihood Ratio Test

A third way to test the assumption of unidimensionality is to check whether variables other than the intended latent dimension, observable teaching skill, affect the item difficulty parameters. This is also important, because the observation instrument must be suitable for use with teachers who have different characteristics like gender and teaching experience, or work with different subject matters or different class sizes. We used Andersen’s (1973, 1977) log-likelihood ratio test that is implemented in the eRm R-package (Mair & Hatzinger, 2007) to compare the difficulty parameters b for each item and to compute Anderson’s log-likelihood ratio χ^2 test. Results are shown in Table 8.4.

Andersen’s log-likelihood ratio test results showed that the difficulty parameters of

- male and female teachers,
- beginning and experienced teachers,
- teachers teaching beta-subject matter (math, science, information science and so on) on the one side and alfa and gamma subjects such as language, English and other subject matters on the other side,
- teachers working in small or large classes were invariant.

When we apply a general norm of .05 for the p-value of Andersen's log-likelihood ratio test, we found a small incident with item 27: "The teacher teaches students how to simplify complex problems". This item has a bit different item difficulty for female and male teachers.

4.4.2 Local Stochastic Independence

Local stochastic independence is one of the underlying assumptions of the Rasch model. The variable measured with a Rasch scale explains why the observed items are related to another. This assumption means that the observed items of a Rasch scale are **conditionally independent** of each other given the score on the **latent variable** that is measured by the Rasch scale. The assumption of local stochastic independence involves that the correlations between the items disappear when the effect of the intended latent variable (teaching skill) has been partialled out. We will use one overall procedure to test whether the 32 items meet this assumption and two item-specific procedures to detect the item pairs susceptible to local dependency.

Confirmatory Factor Analysis with all Residual Correlations Fixed at 0

Firstly we used confirmatory factor analysis (with the Mplus 7.4 program) to check the item correlations after the effect of the latent skill was partialled out. We formulated a one-factor model in which all residual correlations were set at zero.

Table 8.5 shows that both the Comparative Fit Index (CFI) and the Tucker-Lewis index (TLI) are above .90 and the Root Mean Square Error of Approximation (RMSEA) is below the norm of .08, which can be interpreted as an overall indication of local stochastic independence.

Computing Correlations Between the Residues of 32 Items

Using the Mplus 7.4 program, we computed (for the one-factor-model with free residual correlations) the residual correlations of the pairs of items after the effect of the intended latent variable (teaching skill) has been partialled out.

It turned out that 354 out of 496 residual correlations were below .10. A total of 141 residual correlations were between .10 and .30. Only one residual correlation was above .30. The residual correlation between item 22 (The teacher clearly

Table 8.5 Confirmatory factor analyses on 32 dichotomous items and 1 factor residual correlations set at 0

Model fit for residual correlations set at 0	CFI	TLI	RMSEA
Norm	>.90	>.90	<.08
Result	.945	.945	.057

specifies the lesson aims at the start of the lesson) and item 23 (The teacher evaluates whether the lesson aims have been reached) was .318. The residual correlation between item 22 and item 23 goes together with an R squared of .101.

Cohen (1988) evaluates an R below .10 as negligible and an R between .10 and .30 as a small effect. With the exception of the residual correlation between item 22 and item 23, these results might be interpreted as an indication of the local independence of the items.

Chen and Thissen's $LD\chi^2$ Index

Chen and Thissen (1997) proposed a standardized index, the $LD\chi^2$ index, to establish whether there is a violation of the assumption of local stochastic independence for pairs of items. A value of <5 means that there is little likelihood of local dependence. Values between 5 and 10 form a "grey area". When the Chen-Thissen $LD\chi^2$ has a value >10 , it indicates possible local dependence. We computed Chen-Thissen's $LD\chi^2$ with the program IRTPRO (Cai et al., 2005–2013). Results show that some pairs of items indicate possible local dependence ($LD\chi^2 > 10$):

- $LD\chi^2:10.1$: item 2 "maintains a relaxed atmosphere" with item15 "gives a clear explanation of how to use didactic aids and how to carry out assignments"
- $LD\chi^2:12.1$: item 5 "ensures the lesson proceeds in an orderly manner" with item18 "stimulates learners to think about solutions"
- $LD\chi^2:10.4$: item 9 "presents and explains the subject material in a clear manner" with item 24 "offers weaker learners extra study and instruction time"
- $LD\chi^2:10.6$: item14 "teaches in a well-structured manner" with item 17 "stimulates the building of self-confidence in weaker learners"
- $LD\chi^2:11.5$: item 22 "clearly specifies the lesson aims at the start of the lesson" with item 23 "evaluates whether the lesson aims have been reached".

According to this index, we have five pairs of items with possible local dependence. Only the relatively high $LD\chi^2:11.5$ of the last pair of items (22/23) is in agreement with the actual correlation (.318) we have computed between the residuals of these items.

4.4.3 Parallelism of Item Characteristic Curves

Within the Rasch model, the probability of a positive score on an item should depend on the ability of a person, in our case the teacher. When the probability of a positive score on an item is plotted against the skill of teachers, the result would be a smooth S-shaped curve, called the item characteristic curve. The items in the scale should have a stable sequence for each ability group. This means that the item characteristic curves of the items should ideally be parallel. Examining whether certain items have too flat item or too steep characteristic curves, is important, because

these items function differently for people with different skills. We used various procedures to check whether this was the case for the 32 items in the scale.

Anderson's Log Likelihood Ratio Test for Teachers with Low and High Scores

Firstly, we used Andersen's (1973, 1977) log-likelihood ratio test to examine the equality of the item parameters of teachers with a high and low skill level. We used the eRm R-package (Mair & Hatzinger, 2007) to compare the difficulty parameters (b) for each item and to compute Anderson's log-likelihood ratio χ^2 test. Results are shown in Table 8.6.

Results show that with all 32 items Anderson's log likelihood ratio χ^2 test is 74.25 with 31 degrees of freedom and a p-value of .000, indicating a misfit. Leaving out item 17, 20, 31 show that the χ^2 is relatively small, given the number of degrees of freedom (28). The p-value is now .08, also indicating a reasonable fit. The misfitting items are: "item 17, stimulates the building of self-confidence in weaker students", "item 20, let students think aloud", and "item 31, encourages students to think critically". Following this test results, the other 28 items should have about the same difficulty parameters for teachers with a high and a low level of teaching skill. This is a first indication of parallelism of these 28 item characteristic curves.

The Slopes of the Item Characteristic Curves

Another way for testing parallelism is computing the actual slope of each item characteristic curve. We used the LTM R-package (Rizopoulos, 2006) for estimating the slope of the item characteristic curve of each item. The slopes and their standard errors are found in Table 8.7.

The average slope (also called as a parameter in the IRT terms) is 2.01. The rule of thumb for parallelism of item characteristics curves may be that a deviation of approximately two standard errors is too large. Slope parameters that are more than about two times their standard error (S.E.) higher than the average slope parameter are too steep. Slope parameters that are more than about two times their standard error (S.E.) smaller than the average slope parameter are too flat.

The slope of item 9 ("presents and explains the subject material in a clear manner") is rather steep (3.17). The slopes of item 20, 22, and 31 are rather flat. These

Table 8.6 Anderson's log likelihood ratio test for teachers with low and high scores

	Anderson ICC χ^2	df	p-value
32 items	74.246	31	.000
31 items, excluding item 20	60.864	30	.001
30 items, excluding item 20 and 31	43.913	29	.037
29 items, excluding item 20, 31 and 17	39.388	28	.075

Table 8.7 Slopes of the item characteristic curves

Item	Slope (a)	s.e.
1	1.6675	.2613
2	1.5649	.2549
3	1.7911	.2404
4	2.0916	.2506
5	1.6913	.2648
6	2.7299	.3595
7	1.8832	.2504
8	1.5001	.2323
9	3.1681	.4919
10	1.7026	.2284
11	2.7722	.3663
12	2.8507	.3694
13	1.5471	.2094
14	2.5233	.3236
15	1.9324	.2454
16	1.7006	.2201
17	1.5155	.1800
18	2.5843	.3004
19	2.2616	.2672
20	1.1702	.1740
21	1.7998	.2503
22	1.2960	.2066
23	1.8420	.2252
24	2.2747	.2686
25	2.5902	.3003
26	2.4196	.2770
27	1.7772	.2098
28	1.8963	.2351
29	2.3890	.2774
30	1.7950	.2296
31	1.2427	.1637
32	2.2183	.2510

items are respectively “let students think aloud”, “clearly specifies the lesson aims at the start of the lesson”, “encourages students to think critically”.

4.4.4 Conclusions About the Fit of the Rasch Model

At the moment there is no simple approach to test whether a dataset satisfies the assumptions of the Rasch model. Therefore, we have used several different procedures, implemented in several different statistical packages. The use of many

procedures brings along that always one or more items give significant misfit. Some items however, produced several times a misfit:

- Item 9 “presents and explains the subject material in a clear manner” had a too high $LD\chi^2$ (10.4) with item 24 and had a slope of the item characteristic curve that was too steep (3.17).
- Item 20 “lets learners think aloud” disturbed the parallelism of the item characteristic curves with both a significant result on the Andersen’s log likelihood ratio test for high and low scorers, and a too flat slope parameter (1.70).
- Item 22 “clearly specifies the lesson aims at the start of the lesson” showed a too high residual correlation (.318) with item 23, a too high $LD\chi^2$ (11.5) with item 23, a too flat slope parameter (1.30), and a significant result on the Andersen’s log likelihood ratio test for high and low scorers.
- Item 31 “encourage learners to think critically” had significant result on the Andersen’s log likelihood ratio test for high and low scorers and a too flat slope parameter (1.24).

These four items will bring along some problems in determining the zone of proximal development of individual teachers. Therefore, we will remove item 9, 20, 22 and 31 from the scale.

4.5 *The Person Fit*

Thus far, attention was given to items that disturb the fit of the Rasch model. Now the person fit is considered. There are persons having unexpected item score patterns, that should not be expected when the data fit the Rasch model. In the deterministic Guttman model, persons should not respond correctly to difficult items when they respond wrongly to easier items. In the Rasch model, this requirement is somewhat more relaxed, but the number of Guttman errors should remain within certain limits. This is especially true when we want to use a person’s score to detect a person’s zone of proximal development. Several statistics are used to test a person’s fit (Mousavi et al., 2016). In this study, we will use the G-normed-statistic (Meijer, 1994).

4.5.1 Meijer’s G-Normed-Index

The simple G-statistic counts the number of (0, 1) pairs given that the items are ordered in decreasing proportion-correct scores order. The size of the G-statistic depends on the amount of (pairs of) items. The G-normed-statistic was created to bind the G-statistic between zero and one by dividing it by its maximum (Van der Flier, 1982; Meijer, 1994; Tendeiro, 2014). We used the Per Fit R-package (Mousavi et al., 2016) to compute the G-normed-statistic for each observed teacher. Table 8.8

Table 8.8 Meijer's G normed index (average: .21; standard deviation: .18)

G normed index	<.30	.30–.50	>.50
% of observed teachers	72.7	21.4	5.9

presents the results. In an empirical study of Van der Lans et al. (2016) the norm of .30 is proposed for this person fit index.

In 5.9% of the cases the G-normed-index is above 50%, 21.4% of the observed teachers have a G-normed-index between .30 and .50, and 72.7% of the teachers have a G-normed index of <.30.

In the existing statistical literature, we did not find a norm for the G-normed-statistic yet. If we accept the proposal of Van der Lans et al. (2016), a GFI of .30 and more seems too high to be used as a cut-off. This means that we should be careful to use the results for finding a person's zone of proximal development in about 27% of the cases.

Most of these teachers with a high (>.30) G-normed-index are found by four observers who observed each around 20 teachers and by three other observers who observed just one or two teachers. These seven observers have on average five years less experience as a teacher than the other observers do. This difference is significant ($p = .000$). To avoid that this difference affects the result significantly, it is important that these teachers were observed (several) more times, before we could estimate their zone of proximal development more precisely. Another, perhaps simpler approach could be to develop a variant of the G-normed index that can be used in the training of observers. It is also important that observers themselves have sufficient experience in teaching. In the future it might be important to exclude novice teachers from acting as observers in research.

5 Results

Based on results above, we found that the ICALT observation scale with 28 items fulfil the criteria of the Rasch model. In the next part of this chapter, we will present the items, their difficulty parameters and the person parameters of each observed teacher.

5.1 Item Difficulties and Person Parameters

We used the eRm R-package (Mair & Hatzinger, 2007) to compute the difficulty parameter b for each of the dichotomized 28 selected items. Table 8.9 shows our version of a slightly changed Wright map. In column, two and three the items are presented in the order of their difficulty parameter (b) with their standard errors (S.E.).

Table 8.9 Wright map for the ICALT28-scale (N = 375 Korean secondary school teachers)

Item	b	se	Warm's θ	se	Frequency	Cumulative frequency
			-4.477	1.480	1.1	1.1
			-3.313	.878	.3	1.3
			-2.736	.700	.5	1.9
			-2.331	.608	1.1	2.9
			-2.011	.551	1.6	4.5
			-1.740	.512	2.9	7.5
Maintains a relaxed atmosphere	-1.656	.163				
Ensures the lesson proceeds in an orderly manner	-1.549	.159				
			-1.501	.484	2.7	10.1
Shows respect for students in his/her behaviour and language	-1.447	.156				
Uses the time for learning efficiently	-1.348	.153				
			-1.284	.463	3.2	13.3
			-1.083	.448	3.5	16.8
			-.894	.436	1.9	18.7
Gives interactive instructions	-.837	.141				
			-.713	.428	2.4	21.1
Promotes students' self-confidence	-.656	.139				
Provides effective classroom management	-.636	.138				
Presents and explains the subject material in a clear manner	-.557	.137				
			-.538	.421	2.4	23.5
Encourages students to do their best	-.519	.137				
Monitors to ensure students carry out activities in the appropriate manner	-.442	.136				
Teaches in a well-structured manner	-.423	.136				
			-.367	.417	4.0	27.5
Engages all students in the lesson	-.348	.135				
Stimulates the application of what has been learned	-.310	.135				
Offers activities and work forms that stimulate students to take an active approach	-.291	.134				
			-.198	.415	4.3	31.7
Gives a clear explanation of how to use didactic aids and how to carry out assignments	-.089	.133				
			-.030	.415	4.3	36.0

(continued)

Table 8.9 (continued)

Item	b	se	Warm's θ	se	Frequency	Cumulative frequency
Stimulates the use of control activities	.002	.133				
During the presentation stage, checks whether students have understood the subject material	.056	.132				
			.138	.417	4.8	40.8
Evaluates whether the lesson aims have been reached	.182	.132				
Fosters mutual respect	.218	.132				
			.309	.420	3.5	44.3
Asks questions which stimulate students to reflect	.470	.132				
			.483	.426	4.0	48.3
Teaches students to check solutions	.632	.133				
			.663	.433	3.5	51.7
Stimulates students to think about solutions	.723	.133				
			.850	.444	4.3	56.0
Teaches students how to simplify complex problems	.852	.134				
			1.048	.457	3.2	59.2
			1.259	.474	2.9	62.1
Adjusts the processing of subject matter to relevant inter-student differences	1.211	.138				
Adjusts instruction to relevant inter-student differences	1.309	.139				
Asks students to reflect on practical strategies	1.369	.140				
Stimulates the building of self-confidence in weaker students	1.369	.140				
			1.488	.497	2.9	65.1
			1.742	.527	4.0	69.1
			2.032	.568	3.2	72.3
			2.376	.629	3.2	75.5
Offers weaker students extra study and instruction time	2.716	.170				
			2.813	.725	5.6	81.1
			3.434	.909	4.8	85.9
			4.659	1.525	14.1	100.0

The item sequence is more or less similar to the item sequence found in previous studies with Dutch teachers in secondary education (Van de Grift et al., 2014; Van der Lans et al., 2016, 2017). The easiest items are the items about a safe learning climate and efficient classroom management. These items are followed in difficulty with items about the quality of basic instruction. Next items on the dimension are about activating students, teaching learning strategies, and the dimension end with differentiation of teaching, which are the most difficult ones. We will use this ordering in categories of items as indications of the zones of proximal development.

There is one important exception in this ordering. In the previous Dutch study, the item ‘fosters mutual respect’ has a difficulty parameter that is much lower than in the current Korean study (cf. Van de Grift et al., 2014; Van der Lans et al., 2016, 2017).

The person parameters were estimated using Warm’s weighted likelihood estimates (Warm, 1989). This procedure is less biased in comparison with the traditional maximum likelihood estimates method (Hojtink & Boomsma, 1995) and has the advantage that it also can be used to estimate the skills of people with a zero and a maximum score. We used the program WINMIRA (Von Davier, 1994) to compute the person parameters Warm’s weighted likelihood estimates. Table 8.9 shows in column four and five the Warm’s θ and the standard error and some information on the frequency distribution is found in column six and seven.

5.2 *Warm’s θ and some Teacher, Class and School Characteristics*

Table 8.10 presents some descriptive information about the characteristics of the frequency distribution of Warm’s θ .

The average score is 1.03 with a standard deviation of 2.09. Both skewness and kurtosis are <1.0 , which is in indication for an approximately normal distribution. Nevertheless we can observe in Table 8.11 that the amount of teachers with a perfect score ($\theta = 4.66$) is rather high (14%).

Table 8.11 presents some details about relationships of teachers, classrooms and schools and the skill of teachers. We found no significant differences between male and female teachers, teachers teaching α - γ - and β -subject matters or teachers working in general and vocational schools, or working in public or private schools. There was no significant relationship between the years of experience of a teacher and teaching skill. We found a significant, but small, negative correlation of -0.25 between class size and the skill shown by teachers: Teachers show lower skill in large classrooms. Furthermore, we found a significant difference between the skill of teachers in lower and upper secondary education. The difference is 55% of a standard deviation in the advantage of the teacher in lower secondary education.

Table 8.10 Relations between teacher and school characteristics and Warm's θ

	n	average	standard deviation	effect size	significant	R with θ	significant
Theta-score of all 28 ICALT-items	375	1.03	2.09				
Male	183	.95	2.27	.077	.470		
Female	192	1.11	1.90				
Years of experience	369					.095	.069
Subject α - γ	245	1.06	2.08	.033	.700		
Subject β	130	.98	2.11				
Class size	351					-.246	.000
Lower secondary	154	1.69	2.19	.551	.000		
Upper secondary	221	.58	1.88				
General	361	1.04	2.11	.053	.852		
Vocational	14	.93	1.15				
Public	223	.91	1.76	.135	.213		
Private	151	1.19	2.48				
Student's academic engagement	375	3.10	.69			.68	.000

Table 8.11 Areas of proximal development

Zone	Warm's θ			Description	% lessons
1			<-1.0	Safe climate and efficient classroom management	16.8
2	-1.0	-	0.0	Basic tasks of teaching and activating students	19.2
3	0.0	-	1.0	Teaching how to learn	20.0
4	1.0	-	3.0	Differentiating teaching	25.1
5	3.0		4.00	Satisfies the basic and (almost all) advanced teaching skills	4.8
6			>4.0	Satisfies all teaching skills	14.1
					100.0

5.3 Predictive Value of the Scale

In order to study the predictive validity of the Rasch scale we developed a simple scale for measuring the students' academic engagement.

The scale consists of three items that reflect increasing student involvement: 'the learners are fully engaged in the lesson', 'the learners show that they are interested' and 'the learners take an active approach to learning'. The students' academic engagement scale has a range of 1–4. We found an average score of 3.10 with a standard deviation of .69 (cf. Table 8.10). The theta-score of the 28-ICALT-scale had a correlation of .68 with the students' academic engagement scale. So the better the teaching skill, the better the students were involved in the lesson. This is an indication of the predictive validity of the ICALT28-scale.

5.4 *A Proposal for Detecting a person's Zone of Proximal Development*

The raw score of a perfect Guttman scale predicts which items are responded correctly or not. This is very helpful and very precise for finding a person's zone of proximal development. The stochastic character of a Rasch scale, however, brings along several uncertainties in finding a person's zone of proximal development. We have already seen in Table 8.8 that 27% of the observed teachers have severe deviations from the perfect Guttman model. But even when the items have Q-indices (Rost & Von Davier, 1994) nicely near zero and when we wait for more observations for persons with high G-normed-indices (Meijer, 1994), we still have concerns with finding the exact zone proximal development of the observed teachers. The reasons for these concerns are found in the stochastic character of a Rasch scale. Therefore, we will propose an overall procedure with areas of proximal development, based on the meaning of the items. In order to reduce uncertainties in finding a person's zone of proximal development we will use 'areas of proximal development', instead of separate items.

The easiest items are the items about safe learning climate and efficient classroom management. These sets of items are followed in difficulty with a group of items about the quality of basic instruction. Items that are more difficult are about activating students, teaching learning strategies, and the group of items about differentiation of teaching, are the most difficult ones. Inspecting Table 8.9 makes clear that more or less the same ordering is found in the Rasch scale. We will use this ordering in domains of items as indications of the zones of proximal development. Our proposal is laid down in Table 8.11.

Next sections give some descriptions of these areas of proximal development. The scores are clustered in six categories. We used the Warm's θ scores: below -1 ; $-1-0$; $0-1$; $1-3$; $3-4$; and above 4. These are all intervals of just one interval point on the Warm's θ scale. Only one interval is larger ($1-3$) larger. This had to do with the most difficult item. This is of course an arbitrary format, but it guarantees a simple application. The meaning of the categories is just the concept that fits with the meaning of the items within each category. The meaning of the categories corresponds with the complexity level of the teaching skill ranging from low complexity to high complexity. We will present the percentage of lessons we found for each domain.

5.4.1 **Safe Climate and Efficient Classroom Management**

In 16.8% of the observed lessons, the θ -score is below -1.0 . In these lessons, creating a safe learning climate and in maintaining an orderly classroom management was not sufficient. E.g., the atmosphere in the classroom is not relaxed, the lesson does not proceed in an orderly manner and the time for learning is not used efficiently. When there were no special events during the lesson or special other reason

for this low score, than it is clear that the zone of proximal development of teachers within this group is working on a safe climate and an orderly classroom management.

5.4.2 Basic Tasks of Teaching and Activating Students

In 19.2% of the lessons, the θ -score lies between -1.0 and 0.0 . These lessons could be improved by e.g. giving more structured and more interactive instructions.

5.4.3 Teaching Students How to Learn

In 20.0% of the lessons, the θ -score is between 0.0 and 1.0 . In these lessons, the basic skills of teaching (creating a safe and stimulating educational climate, an orderly classroom management, and clear and activating instruction) are sufficient.

These lessons could be improved by teaching students how they can learn things: The teacher can improve the lesson by e.g. asking questions that stimulate students to reflect and to check solutions.

5.4.4 Differentiating Teaching

In 25.1% of the lessons, the basic tasks of teaching, activating students, and teaching students how to learn things are observed to be sufficient. These lessons have θ -scores between 1.0 and 3.0 . These lessons can be improved by adjusting instruction and the processing of subject matter to relevant inter-student differences. One of the most difficult tasks for the teachers in this zone of proximal development is offering weaker students extra study and instruction time.

5.4.5 Lessons Satisfying All Basic and Almost All Advanced Teaching Skills

In 4.8% of the lessons, a θ -score between 3.0 and 4.0 is found. Teachers reveal in these lessons all basic skills and most advanced teaching skills.

5.4.6 Lessons Satisfying all Teaching Skills

In 14.1% of the lessons, all 28 teaching skills were exhibited. This is a rather high percentage. The percentage of 14% perfect scores could be a reason to add some more important items with higher difficulty to this scale. We know that the current

version of the ICALT observation instrument can be supplemented with additional items about differentiation.

These somewhat arbitrary areas are mostly important for giving a θ -score a meaning in terms of the skills of teachers. The θ -score is the actual level of development, and the domain (cf. Table 8.9) specifies the zone of proximal development. The limits used for these domains are of course somewhat arbitrary. When a lesson gets a score that is just below the upper limit of one of the different domains, it is probably wise to shift the zone of proximal development to the next area. To give an example: A teacher with a score of Warm's $\theta = .85$ (cf. Table 8.9) does not really have to wait until he masters the last item of teaching how to learn, before he can start differentiation of his instruction.

6 Conclusions

In this study, we reported the development of a 28-item-scale for observing teaching skills that fulfils the assumptions of the dichotomous Rasch model.

We discovered that the order of item difficulty found among Dutch secondary school teachers is in general maintained among secondary school teachers from a totally different culture, the South Korean culture. There is one important exception in this ordering. In the previous Dutch study, the item 'fosters mutual respect' has a difficulty parameter that is much lower than in the current Korean study. This is probably due to the fact that the word 'respect' in Asian cultures has a more stringent meaning than in many Western European cultures. This makes it necessary to conduct further and more detailed research into cultural differences in the quality of teaching skill.

The scores on the scale had predictive value for the engagement of students. In subsequent studies it should be determined whether the scale also has a predictive value for the performance of the students.

With this study, we have developed an observation tool with which we can not only determine the current level of development of a teacher, but we also can give an indication of the zone of proximal development of the observed teacher. The latter in particular is very important. It simply does not help enough if we tell a teacher what his or her score is and what s/he does not do well. The 'trick' is to help a teacher by pointing out activities that s/he does not do, but that are within her or his reach. This ICALT observation instrument offers the possibility to coach teachers and guide them in matters that they are not yet doing.

Appendix

ICALT Lesson Observation Form (international comparison of learning and teaching)	
Country	...
School name	...
Location	...
Level of education	0=primary education 1=secondary education
School denomination	0=public 1=private
Subject matter	...
Name teacher	...
Gender teacher	M / F
Years of teaching experience teacher	...
Observe the following behaviours and events:	
Rate! Please circle the appropriate answer:	
1= mostly weak; 2=more often weak than strong; 3= more often strong than weak; 4= mostly strong	
Observed? Please circle the appropriate answer:	
0= no, I have not observed this; 1= yes, I have observed this	
Date observation (dd-mm-yyyy)	...
Class
No of learners present
Name observer
Gender observer	M / F
Years of teaching experience observer	...
Occupation observer	0=school teacher 1=university teacher 2=other Please specify...

Domain	Indicator: The teacher...	Rate ¹	Examples of good practice: The teacher...	Observed ²
Safe and stimulating learning climate	1 ...shows respect for learners in his/her behaviour and language	1 2 3 4	...lets learners finish their sentences	0 1
			...listens to what learners have to say	0 1
			...does not make role stereotyping remarks	0 1
	2 ...maintains a relaxed atmosphere	1 2 3 4	...addresses learners in a positive manner	0 1
			...uses and stimulates humour	0 1
			...accepts the fact that learners make mistakes	0 1
	3 ...promotes learners' self-confidence	1 2 3 4	...shows compassion and empathy for all learners present	0 1
			...gives positive feedback on questions and remarks from learners	0 1
			...compliments learners on their work	0 1
	4 ...fosters mutual respect	1 2 3 4	...acknowledges the contributions that learners make	0 1
			...stimulates learners to listen to each other	0 1
			...intervenes when learners make fun of someone	0 1
			...keeps (cultural) differences and idiosyncrasies in mind	0 1
			...stimulates solidarity between learners	0 1
			...encourages learners to experience activities as group events	0 1

(continued)

ICALT Lesson Observation Form (international comparison of learning and teaching)			
Efficient organi-sation	5 ...ensures the lesson proceeds in an orderly manner	1 2 3 4 <i>Learners enter and settle in an orderly manner</i>	0 1
		<i>...intervenes timely and appropriately in case of disorder</i>	0 1
		<i>...safeguards the agreed rules and codes of conduct</i>	0 1
		<i>...keeps all learners involved in activities until the end of the lesson</i>	0 1
		<i>...makes sure that learners know what to do if they need help with their work and explains clearly when they can ask for help</i>	0 1
		<i>...makes sure learners know what to do when they have finished their work</i>	0 1
	6 ...monitors to ensure learners carry out activities in the appropriate manner	1 2 3 4 <i>...checks whether learners have understood what they have to do</i>	0 1
		<i>...provides feedback on learners' social functioning whilst carrying out a task</i>	0 1
	7 ...provides effective classroom management	1 2 3 4 <i>...explains clearly which materials can be used</i>	0 1
		<i>The materials for the lesson are ready for use</i>	0 1
		<i>Materials are geared at the right level and developmental stage of the learners</i>	0 1
	8 ...uses the time for learning efficiently	1 2 3 4 <i>... starts the lesson on time</i>	0 1
<i>... does not waste time at the beginning, during, or at the end of the lesson</i>		0 1	
<i>...prevents any unnecessary breaks from occurring</i>		0 1	
<i>...does not keep learners waiting</i>		0 1	

Clear and structured instructions	9 ...presents and explains the subject material in a clear manner	1 2 3 4	...activates prior knowledge of learners	0	1
			...gives staged instructions	0	1
			...poses questions which learners can understand	0	1
			...summarises the subject material from time to time	0	1
	10 ...gives feedback to learners	1 2 3 4	...makes clear whether an answer is right or wrong	0	1
			...makes clear why an answer is right or wrong	0	1
			...gives feedback on the way in which learners have arrived at their answer	0	1
			...creates learners assignments which stimulate active participation	0	1
	11 ...engages all learners in the lesson	1 2 3 4	...asks questions which stimulate learners to reflect	0	1
			...makes sure that learners listen and/or continue working	0	1
			...allows for 'thinking time', after asking a question	0	1
			...also invites learners to participate who do not volunteer to do so	0	1
12 ...during the presentation stage, checks whether learners have understood the subject material	1 2 3 4	...ask questions which stimulate learners to reflect	0	1	
		...checks regularly whether learners understand what the lesson is about	0	1	
		...praises learners who do their best	0	1	
		...makes clear that all learners should do their best	0	1	
13 ...encourages learners to do their best	1 2 3 4	...expresses positive expectations about what learners are going to achieve	0	1	

(continued)

ICALT Lesson Observation Form (international comparison of learning and teaching)				
14	...teaches in a well-structured manner	1 2 3 4	The lesson is built up in terms of clear stages and transitions between stages	0 1
			The lesson builds up logically, going from the simple to the complex	0 1
			Activities and assignments are connected to the materials presented during the presentation stage	0 1
15	...gives a clear explanation of how to use didactic aids and how to carry out assignments		The lesson offers a good variety of presentation, instruction, controlled practice, free practice, and so forth.	0 1
		1 2 3 4	...makes sure that all learners know what to do	0 1
			...explains how lesson aims and assignments relate to each other	0 1
16	...offers activities and work forms that stimulate learners to take an active approach		...explains clearly which materials and sources can be used	0 1
		1 2 3 4	...uses diverse forms of conversation and discussion	0 1
			...offers controlled (pre-)practice	0 1
Intensive and activating teaching			...lets learners work in Group	0 1
			...uses Information and Communication Technology (ICT, e.g., digiboard, beamer)	0 1
			...employs a variety of instruction strategies	0 1
			...varies assignments	0 1
			...varies lesson material	0 1
			...uses materials and examples from daily Life	0 1
			...asks a range of questions	0 1
		1 2 3 4	...gives positive feedback on questions from weaker learners	0 1
			...displays positive expectations about what weaker learners have to achieve	0 1
			...compliments weaker learners on their works	0 1
	...acknowledges the contributions made by weaker learners	0 1		

ICALT Lesson Observation Form (international comparison of learning and teaching)			
Adjusting instructions and learner processing to inter-learner differences	23	...evaluates whether the lesson aims have been reached	0 1
		...evaluates learners' performance	0 1
	24	...offers weaker learners extra study and instruction time	0 1
		...gives weaker learners extra instruction Time	0 1
		...gives weaker learners extra exercises/practices	0 1
Teaching learning strategies	25	...adjusts instructions to relevant inter-learner differences	0 1
		...puts learners who need little instructions (already) to work	0 1
		...gives additional instructions to small groups or individual learners	0 1
		...does not simply focus on the average learner	0 1
		...distinguishes between learners in terms of the length and size of assignments	0 1
Teaching learning strategies	26	...adjusts the processing of subject matter to relevant inter-learner differences	0 1
		...allows for flexibility in the time learners get to complete assignments	0 1
		...lets some learners use additional aids and means	0 1
		...teaches learners how to simplify complex problem	0 1
		...teaches learners how to break down complex problems into simpler Jones	0 1
Teaching learning strategies	27	...teaches learners how to simplify complex problems	0 1
		...teaches learners to order complex problem	0 1
	28	...stimulates the use of control activities	0 1
		...pays attention to prediction strategies for reading	0 1
		...lets learners relate solutions to the context of a problem	0 1
Teaching learning strategies	29	...teaches learners to check solutions	0 1
		...stimulates the application of alternative strategies	0 1
		...teaches learners how to estimate outcomes	0 1
		...teaches learners how to predict outcomes	0 1
		...teaches learners how to relate outcomes to the practical context	0 1

30	...stimulates the application of what has been learned	1 2 3 4	... stimulates the conscious application of what has been learned in other (different) learning contexts ...explains to learners how solutions can be applied in different situations ...relates problems to previously solved problem ...asks learners to provide explanations for occurrences ...asks learners for their opinion ...asks learners to reflect on solutions or answers given ...asks learners to provide examples of their own ...asks learners to explain the different steps of the strategy applied ...gives an explicit explanation of possible (problem-solving) strategies ...asks learners to expand on the pros and cons of different strategies	0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1
31	...encourages learners to think critically	1 2 3 4	Examples of good practice: Learners...	Observed ²
32	...asks learners to reflect on practical strategies	1 2 3 4	...pay attention during instructions are given ...participate actively in conversations and discussions ...ask questions ...listen actively when instructions are being given ...show their interest by asking follow-up questions ...ask follow-up questions ...show that they take responsibility for their own learning process ...work independently ...take the initiative themselves ...use their time efficiently	0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1
Indicator: The learners...	Rate ¹			
Learner engagement	33 ...are fully engaged in the lesson	1 2 3 4		
	34 ...show that they are interested	1 2 3 4		
	35 ...take an active approach to learning	1 2 3 4		

References

- Andersen, E. B. (1973). A goodness of fit test for the Rasch model. *Psychometrika*, *38*, 123–140.
- Andersen, E. B. (1977). Sufficient statistics and latent trait models. *Psychometrika*, *42*, 69–81.
- Bell, C. A., Dobbelaer, M. J., Klette, K., & Visscher, A. (2018). Qualities of classroom observation systems. *School Effectiveness and School Improvement*, *30*, 3–29. <https://doi.org/10.1080/009243453.2018.1539014>
- Cai, L., Thissen, D., & Du Toit, S. (2005–2013). *IRTPRO 2.1*. Scientific Software International.
- Chen, W.-H., & Thissen, D. (1997). Local Dependence Indexes for Item Pairs Using Item Response Theory. *Journal of Educational and Behavioral Statistics* *22*(3), 265–289.
- Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods & Research*, *36*(4), 462–494.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modelling*, *9*(2), 233–255.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290.
- Cohen, J. A. (1960). Coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.
- Cohen, J. (1967). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- De Jager, T., Coetzee, M.J., Maulana, R., Helms-Lorenz, M., & Van de Grift, W. (2017). Profile of South African secondary-school teachers' teaching quality: Evaluation of teaching practices using an observation instrument. *Educational Studies*, *43*(4), 410–429.
- DeMars, C. (2010). *Item response theory*. Oxford University Press.
- Dobbelaer, M. J. (2019). *The quality and qualities of classroom observation systems*. Ipskamp printing.
- Falotico, R., & Quatto, P. (2015). Fleiss' kappa statistic without paradoxes. *Quality & Quantity*, *49*, 463–470.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). John Wiley.
- Fleiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement*, *33*, 613–619.
- Guttman, L. A. (1950). The basis for scalogram analysis. In S. A. Stouffer, Guttman, L. A., & E. A. Schuman Measurement and prediction. Volume 4 of studies in social psychology in world war II. : Princeton University Press
- Hallgren, K. V. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23–34.
- Hanushek, E. A., & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review: Papers & Proceedings*, *100*(2), 267–271.
- Hattie, J. A. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Hattie, J. A. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.
- Hojtink, H., & Boomsma, A. (1995). On person parameter estimation in the dichotomous Rasch model. In G. H. Fischer & I. W. Molenaar (Eds.), *Rasch models: Foundations, recent developments, and applications* (pp. 53–68). Springer.
- Houtveen, A. A. M., & Van de Grift, W. J. C. M. (2007a). Effects of metacognitive strategy instruction and instruction time on reading comprehension. *School Effectiveness and School Improvement*, *18*(2), 173–190.
- Houtveen, A. A. M., & Van de Grift, W. J. C. M. (2007b). Reading instruction for struggling learners. *Journal of Education for Students Placed at Risk*, *12*(4), 405–424.
- Houtveen, A. A. M., Van de Grift, W. J. C. M., & Creemers, B. P. M. (2004). Effective school improvement in mathematics. *School Effectiveness and School Improvement*, *15*(3–4), 337–376.

- Houtveen, A. A. M., Van de Grift, W. J. C. M., & Brokamp, S. K. (2014). Fluent reading in special elementary education. *School Effectiveness and School Improvement*, 25(4), 555–569.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
- Kane, T. J., & Staiger, D. O. (2008). *Estimating teacher impacts on student achievement: An experimental evaluation* (NBER Working Paper No. 14607).
- Kirby, A., Gebiski, V., & Keech, A. C. (2002). Determining the sample size in a clinical trial. *The Medical Journal of Australia*, 177(5), 256–257.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). Guilford.
- Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, 88(4), 547–588.
- Kundel, H. L., & Polansky, M. (2003). Measurement of observer agreement. *Radiology*, 228, 303–308.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Lipsey, M. W. (1990). *Design sensitivity: Statistical power for experimental research*. Sage.
- Lompscher, J., & Rückriem, G. (2002). Editorial. In L. S. Vygotskij (Ed.), *Denken und Sprechen*. Beltz Verlag.
- Mair, P., & Hatzinger, R. (2007). Extended Rasch modelling: The eRm package for the application of IRT models in R. *Journal of Statistical Software*, 20, 1–20.
- Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness of fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, 103, 391–410.
- Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers of overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modelling*, 11, 320–341.
- Marzano, R. J. (2003). *What works in schools: Translating research in action*. ASCD.
- Maulana, R., Helms-Lorenz, M., & Van de Grift, W. (2015). Development and evaluation of a questionnaire measuring pre-service teachers' teaching behaviour: A Rasch modelling approach. *School Effectiveness and School Improvement*, 26, 169–194.
- Maulana, R., Helms-Lorenz, M., & Van de Grift, W. (2016). The role of autonomous motivation for academic engagement of Indonesian secondary school students: A multilevel modelling approach. In R. B. King & A. B. I. Bernardo (Eds.), *The psychology of Asian learners. A festschrift in honor of David Watkins* (pp. 237–252). Springer.
- Meijer, R. R. (1994). The number of Guttman errors as a simple and powerful person-fit statistic. *Applied Psychological Measurement*, 18, 311–314.
- Mousavi, A., Tendeiro, J. N., & Younesi, J. (2016). Person fit assessment using the PerFit package in R. *The Quantitative Methods for Psychology*, 12(3), 232–242.
- Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Muthén and Muthén.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Denmark: Paedagogiske Institut.
- Rasch, G. (1961). On general laws and the meaning of measurement in psychology. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, IV* (pp. 321–334). University of California Press.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools and academic achievement. *Econometrica*, 73(2), 417–458.
- Rizopoulos, D. (2006). ltm: An R-package for latent variable modelling and item response theory analyses. *Journal of Statistical Software*, 17(5), 1–25.
- Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation*, 9(4), 1–11.
- Ten Hove, D., Jorgensen, T. D., & Van der Ark, L. A. (2018). On the usefulness of interrater reliability coefficients. In M. Wiberg, S. A. Culpepper, R. Janssen, J. Gonzáles, & D. Molenaar (Eds.), *Quantitative psychology. The 82nd annual meeting of the psychometric society, Zurich, Switzerland* (pp. 67–76). Springer.
- Tendeiro, J. N. (2014). Package 'PerFit' (published online). In R. Cran (Ed.), *The comprehensive R network*. Retrieved from: <http://cran.r-project.org/web/packages/PerFit/PerFit.pdf>

- Tucker, L., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1), 1–10.
- Van der Flier, H. (1982). Deviant response patterns and comparability of test scores. *Journal of Cross-Cultural Psychology*, 13(3), 267–298.
- Van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and an application of an assessment instrument. *Educational Research*, 49(2), 127–152.
- Van de Grift, W. J. C. M. (2014). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement*, 25(3), 295–311.
- Van de Grift, W. (2021). Het coachen van leraren (1/2) [the coaching of teachers]. *Basisschoolmanagement*, 2, 24–29.
- Van de Grift, W. J. C. M., & Lam, J. F. (1998). Het didactisch handelen in het basisonderwijs [Teaching in primary education]. *Tijdschrift voor Onderwijsresearch*, 23(3), 224–241.
- Van de Grift, W., Helms-Lorenz, M., & Maulana, R. (2014). Teaching skills of student teachers: Calibration of an evaluation instrument and its value in predicting student academic engagement. *Studies in Educational Evaluation*, 43, 150–159.
- Van de Grift, W. J. C. M., Chun, S., Maulana, R., Lee, O., & Helms-Lorenz, M. (2017). Measuring teaching quality and student engagement in South Korea and The Netherlands. *School Effectiveness and School Improvement*, 28(3), 337–349.
- Van de Grift, W. J. C. M., Houtveen, A. A. M., Van den Hurk, H. T. G., & Terpstra, O. (2019). Measuring teaching skills in elementary education using the Rasch model. *School Effectiveness and School Improvement*, 30, 455–486. <https://doi.org/10.1080/09243453.2019.1577743>
- Van den Hurk, H. T. G., Houtveen, A. A. M., & Van de Grift, W. J. C. M. (2016). Fostering effective teaching behaviour through the use of data-feedback. *Teaching and Teacher Education*, 60, 444–451.
- Van der Lans, R. M., Van de Grift, W. J. C. M., Van Veen, K., & Fokkens-Bruinsma, M. (2016). Once is not enough: Establishing reliability criteria for feedback and evaluation decisions based on classroom observations. *Studies in Educational Evaluation*, 50, 88–95.
- Van der Lans, R. M., Van de Grift, W. J. C. M., & Van Veen, K. (2017). Individual differences in teacher development: An exploration of the applicability of a stage model to assess individual teachers. *Learning and Individual Differences*, 58, 46–55.
- Van der Lans, R. M., Van de Grift, W. J. C. M., & Van Veen, K. (2018). Developing an instrument for teacher feedback: Using the Rasch model to explore teachers' development of effective teaching strategies and behaviors. *The Journal of Experimental Education*, 86(2), 247–264.
- Viera, A. J., & Garret, J. M. (2015). Understanding interobserver agreement: The kappa statistic. *Family Medicine*, 37(5), 360–363.
- Von Davier, M. (1994). *WINMIRA 200.1*. IPN.
- Vygotsky, L. S. (1930). *Mind and society*. Harvard University Press.
- Warm, T.A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54, 427–450.



Prof. Dr. Wim van de Grift (1951) is emeritus professor in Educational Sciences at the University of Groningen. He was director of the Teacher Training Institute of the University of Groningen and was scientific advisor of the Inspectorate of Education in the Netherlands. Van de Grift's research is aimed at the development and testing of theories on the professional development of teachers. This research program focusses on the following questions: How do teaching skills develop during the teaching career? Which factors influence the development of teaching skills? What is the influence of the teachings skills of teachers on students' academic engagement and students' achievements?

Van de Grift studied psychology at Utrecht University and obtained in 1987 his doctoral degree at Leiden University with a dissertation on ‘The role of the school leader in educational innovations’.

Work:

1978–1989: University of Amsterdam and Utrecht University.

1989–2016: Inspectorate of Education (Ministry of Education, Culture and Science).

2008–2016: University of Groningen.

2017–now: Director of his own company specialized in observing and coaching teachers.



Emeritus Prof. Okhwa Lee (since 2022 March). Department of Education of Chungbuk National University, South Korea. Prof. Ok-hwa Lee is a specialist in educational technology and a practitioner of teacher education. She has been a pioneer of the e-learning, technology applications in education and educational reform through smart education in Korea. She was a member of the Presidential Educational Reform Committee and the Presidential e-Government Committee of the Republic of Korea, also consulting members for various ministries regarding educational applications of technology. She has rich experiences of international collaborations with the Europe Erasmus mobility with Finland Sweden, Estonia, Netherlands and etc., long history of research collaboration with USA, Australia, Thailand and etc. Recently she collaborated with developing countries through the Korean government ODA (Official Development Assistant) programs to Sudan, Nigeria, Nicaragua, Vietnam, Ethiopia, Cambodia, Myanmar, and etc. Her work through ODA focused on teachers’ capacity development of teaching skills using technology.



Prof. Seyeoung Chun is Professor Emeritus of Education at Chungnam National University, one of the major national universities in Daejeon, Korea. He received his education and Ph.D. from Seoul National University, South Korea, and has been actively engaged in education policy research and has held several key positions such as Secretary of Education to the President and CEO of KERIS. He founded the Smart Education Society in 2013, and has led many projects and initiatives for the paradigm shift of education in the digital era. Since his early career at the Korean National Commission for UNESCO, he has participated in many international cooperation projects and worked for several developing countries such as Nicaragua, Honduras, Cambodia, etc. *Education Miracle in the Republic of Korea* is the latest book to be published as a summary of his academic life.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Part II

Effective Teaching: Insights from Specific Countries

Part II Overview

Several contributions to this volume deepen our conceptual and country specific understanding of manifestations of effective teaching. The six chapters of this part contribute towards widening the scope of understanding with rich descriptions the historical-, policy- and daily demands faced by teachers in different contexts, in relation to effective teaching behaviours. The Indonesian study (Chap. 10) describes general and specific profiles of teachers from 13 provinces that offer underpinnings for future professional development programs, towards improving teaching quality in Indonesia. The study presented in Chap. 11 describes the contextual background of teachers and teaching quality in Mongolia through the lens of educational policies, practices and challenges surrounding the teacher, and by describing how the curriculum sets the stage for teaching behaviours. The historical changes in teacher education in India are described in Chap. 12, setting out to measure the quality of the current learning environment reported by student teachers. A legal, epistemological and empirical approach is reported in Chap. 13 to describe factors influencing teaching effectiveness and student engagement in Spain. In Chap. 14 the relationship between the high level of teaching quality measured in South Korea is discussed in the light of teacher education and the educational policy (in- and out of schools) of South Korea. One study focuses on learning environments in Australia (Chap. 15), using a student questionnaire (SPAQ) to identify exemplary teachers. These exemplary science teachers were found to be thorough in their teaching, giving students enough time to prepare for the assessment, allowing students to choose freely from a variety of assessments and being flexible in teaching and assessment.

Two chapters broaden the conceptual scope of effective teaching by focusing on video-taped lessons using different instruments simultaneously to measure dialogical interactions, ICALT and CETIT dimensions (Chap. 9), and by comparing ICALT dimensions to that of inspiring teaching (Chap. 16).

Chapter 9

Dialogic Interactions in Higher Vocational Learning Environments in Mainland China: Evidence Relating to the Effectiveness of Varied Teaching Strategies and Students' Learning Engagement



Yanmin Zhao, Marc Kleinknecht, and James Ko 

Abstract The study aims to explore students' learning in the vocational classroom learning environment and the teaching practices of vocation-oriented subjects in Chinese higher vocational institutions. Based on sixty lesson observations, four selected videotaped lessons were used to conduct in-depth dialogic interaction analysis of teacher-led (the teacher to students), student-led (students to the teacher), students to students, and students to the course content according to ICALT and CETIT dimensions of effective teaching. Vocational collaborative learning and adaptive instructions were analysed through the in-class activities of the learning processes that students were engaged in within the classroom. Findings suggest that dialogic teaching in classrooms enhanced practical understanding in specialised vocational subjects and students' learning engagement, for example, classroom practices such as small group teaching of vocational skills and lesson activities connected to work-related learning situations. The study also reveals that a built-in flexible teaching arrangement stimulates vocational students' involvement in collaborative learning and promotes interactions between students' classroom-based training activities. The study implies that effective dialogic classroom learning environments should integrate vocational students' career learning and work-based instructions.

Y. Zhao (✉)

The Education University of Hong Kong, Tai Po, Hong Kong
e-mail: zhaoy@eduhk.hk

M. Kleinknecht

Leuphana University of Luneburg, Lüneburg, Germany
e-mail: marc.kleinknecht@leuphana.de

J. Ko

Department of Education Policy and Leadership, The Education University of Hong Kong,
Tai Po, Hong Kong
e-mail: jamesko@eduhk.hk

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_9

207

Keywords Dialogic interactions · Vocational teaching effectiveness · Learning engagement

1 Introduction

Research shows that a career-oriented learning environment can enhance students' development of career competencies and can foster students' participation in practice-based learning and vocation-related activities (Kuijpers et al., 2011). Engaging students in an interactive classroom environment has regularly been discussed in various studies. For example, the online classroom environment of internet-based business courses influences students' learning engagement in working with small groups and developing discussion questions (Arbaugh, 2000). Meta-analyses on teaching effectiveness suggest that the effects of teaching on student learning are diverse and complex regarding the integrated components of learning in different contexts (Seidel & Shavelson, 2007). In the context of interactive classrooms, dialogic pedagogy as an approach increases learner engagement and classroom interactions, which enables teachers to value learners' voices and promotes reflective learning (Lyle, 2008). Empirical studies and theoretical summaries on dialogic teaching and learning, dialogic interactions, and dialogic classroom have shown significant impacts on fostering students' engagement in learning and teaching practice (Granger et al., 2012; Haneda, 2016; Lyle, 2008; Mercer & Littleton, 2007).

Other research suggests that effective teachers adapt their instructions in response to the features of classroom activities and students' reflections on using open tasks (Parsons, 2012). Adaptive expertise described by (Darling-Hammond & Bransford, 2007) emphasises on establishing effective classroom instruction connecting to students' learning performance. Teacher participation in learning activities positively relates to the likelihood of effective teaching (de Vries et al., 2015). Therefore, teacher-led and student-led dialogic classroom interactions encourage students' collaborative learning through adaptive dialogic instructions (Gillies, 2019; Kim & Wilkinson, 2019; Teo, 2016). By making vocational learning environments like workplaces, classroom-based activities emphasise flexible activity-based training platforms to facilitate students' learning engagement (Zhao & Ko, 2020), vocational teachers encourage students to engage in work-based learning and assist the transfer of learning from the classroom to many other situations.

It is reasonable to suggest that the importance of understanding vocation-oriented classroom dialogues helps improve students' learning and teaching practices concerning vocational teaching effectiveness. We apply two observational instruments developed for evaluating effective teaching behaviours and inspiring teaching in the vocation-oriented classroom. By selecting four videotaped lessons from different specialised subjects based on the mean scores of the percentile rank of sixty lessons' distribution, in-depth classroom dialogic analysis was conducted to explore vocational students' learning engagement and teachers' dialogic adaptive instructions through Teacher-led-Student, Student-led-Teachers, Student-Student, and Student-Content interactions. Vocational classroom dialogues were used to analyse the

characteristics of the learning engagement of classroom practice and teaching adaptations in the vocational learning environment.

2 Literature Review

2.1 *Dialogic Interactions and Students' Learning in Classroom Settings*

Teachers' and students' dialogic interactions play a key role in engaging students with classroom dialogue to facilitate the exchange of ideas and opinions. Researchers have pointed out that dialogue makes students more active in sharing ideas and enables active participation in the process of dialogic interactions (Rojas-Drummond et al., 2013; Mercer & Littleton, 2007). However, systematic research by Howe and Abedin (2013) on classroom dialogue indicates that classroom dialogues are mainly teacher-student interactions around traditional information-response-feedback, and pedagogic teaching style is also the major factor in determining the student participation and dialogic patterns of group work activities. Other studies on different forms of dialogue such as student-teacher interactions emphasised students' learning through lectures, textbooks, and classroom activities (Granger et al., 2012), and Gillies (2016, 2019) highlights the teacher's role in dialogic teaching, which can be used to develop students' learning proficiency. In scaffolding children's learning and understanding processes, Rojas-Drummond et al. (2013) analysed the dialogic interactions among teachers and students for comprehending teaching and learning in classroom settings. The combined dialogic interactions have a significant effect on students' learning outcomes in online and blended learning environments (Ekwunife-Orakwue & Teng, 2014). However, other studies suggest that it is a highly demanding task in enhancing student engagement through dialogic inquiries and teachers' awareness of dialogic interactions in the classroom may not be commonly emphasised in the classroom discourse (Kumpulainen & Lipponen, 2010; Nystrand et al., 2003).

Others who have investigated the effect of dialogic interactions on students' thinking and learning include collaborative learning through productive dialogues (Gillies, 2019; Vrikki et al., 2019a), dialogic engagement in small group reading comprehension (Maine & Hofmann, 2016), and dialogic classroom fostering students' engagement in learning (Haneda, 2016). Evidence has emerged from these studies that guiding students to engage constructively with each others' ideas contributes to a deeper understanding of disciplinary knowledge and helps students clarify their thinking with a small group and whole-class discussions. Studies by Haneda et al. (2017), Kim and Wilkinson (2019), Teo (2016), and Rojas-Drummond et al. (2013) highlight the importance of the teacher's role in structuring students' interactions with each other around tasks. According to Alexander's (2017) five principles of classroom dialogue, the characteristics of dialogic interactions and teaching should be: (1) collective – with teachers and students in tasks as a group or

a class; (2) reciprocal – with shared ideas and viewpoints between teachers and students; (3) supportive – students encouraging and helping each other to reach common understandings; (4) cumulative – facilitating students in building on their own ideas and extending them into further understanding and enquiry; (5) purposeful – the teacher’s plan is directed towards particular learning goals (p. 28). Therefore, in promoting student engagement and academic dialogue, Gillies (2019) suggests the importance of structuring collaborative learning where students are taught how to advance an argument during group discussion and provide justifications to support their ideas and stance.

2.2 Dialogic Teaching and Adaptive Instructions

In connection with classroom dialogue, dialogic teaching as a pedagogical approach focuses on various pedagogies that foster classroom talk in a specific discourse practice (Kim & Wilkinson, 2019). Alexander’s (2004, 2017) concept of dialogic teaching requires teachers to organise teacher- or student-led small groups and engage students in teacher- or student-directed discussions. The dialogic interaction in coaching sessions helps teachers understand pedagogical approaches in the strategic use of classroom dialogue to teaching and learning (Haneda et al., 2017). Lyle (2008) addresses the dialogic practice relating to the quality of classroom interaction and the engagement of students’ learning, which draws attention to the features of dialogic teaching and learning in small collaborative groups. In considering dialogic engagement in classroom settings, problems and difficulties are also pointed out in implementing dialogic teaching in the higher-level interactions involving constructive meaning-making and reasoning (Lyle, 2008; Maine & Hofmann, 2016). Hardman (2016) emphasises the high quality of classroom talk between teacher-led and student-led interactions in empowering students to obtain transferable skills and stimulating learning experiences. A dialogic teaching intervention plays a central role in small-group dialogues and discussions (Hardman, 2019; Vrikki et al., 2019a), which suggests that the implementation of a dialogic pedagogy in teaching and learning serves to improve students’ participation, engagement and learning.

Adaptive instruction or individualised instruction is similar to orchestration (Dillenbourg, 1999; Dillenbourg, 2013; Dillenbourg & Tchounikine, 2007) that the teacher monitors the real classroom situation and decides what kinds of adaptations are necessary for students and then performs the individualised adaptations to the classroom. Dillenbourg (2013) refers to “orchestration” as a metaphor to indicate how the teacher acts as a conductor to demonstrate “how a teacher manages, in real-time, multi-layered activities in a multi-constraints context” (p. 485). An adaptation model proposed by Deed et al. (2019) suggests that the adaptive process in a flexible learning environment is complex and non-linear, which illustrates that teachers engage with the idea of space as an influence on teaching practice, and consider the relationship between teaching and learning space, and integrate the interplay

between teaching and learning space. This is consistent with the view that flexible physical space enables greater collaboration in the teaching and learning processes and impacts the interplay between student activities and classroom engagement (Dane, 2016). Although teacher adaptation may include changes in teachers' practical knowledge and its interaction with situated experience and affordances of flexible learning environments for teachers to influence student engagement, our focus is mainly on vocational teachers' practices in terms of adaptive transactions between teacher and context, instructions and students with learning materials within classroom dialogues.

2.3 Varied Teaching Effectiveness and Vocation-Oriented Learning Environments

Teaching effectiveness can be diverse in light of the variety of teaching approaches applied in the different contexts of teaching and learning (Seidel & Shavelson, 2007). We define vocational teaching effectiveness as a set of classroom dialogues concerning the dialogic interactions involved in vocation-oriented teaching, students' collaborative learning, and adaptive instructions on classroom training activities. The classroom learning environment includes not only the physical space for learning but also the intangible classroom climate, which strongly influences students' learning outcomes and competence development (Fraser, 2001). Alfassi (2004) finds that the learner-centered environment promotes higher scores in academic achievement and relatively higher motivation for learning. Vocation-oriented learning environments emphasise on students' learning process, which allows vocational students to reflect on their learning, showcase their vocational skills, and collaborate with peers (Valtonen et al., 2012). In relating to students' collaborative learning, vocational dialogues focusing on career guidance methods play an important role in the relationship between vocational learning environment and students' career competencies, which aims to foster students' career learning in some aspects of the learning environment (Kuijpers et al., 2011).

On the other hand, flexibility in the vocational learning environment has been given emphasis with its flexible classroom settings such as activity-based training platforms, computer-supported workshops, and simulated software for practical training (Zhao & Ko, 2020). A flexible vocational learning environment facilitates students' engagement in the process of training as Dillenbourg (2013) suggests that teachers have the freedom to adjust class activities in order to adapt to students' learning needs. Therefore, the interaction of the collaborative learning activity within its relevant environmental context provides a lens for analyzing learning processes in the changing learning environments that students are engaged in within vocational classrooms. As stated by Kuijpers et al. (2011), a flexible vocational learning environment fosters the development of students' career competencies, while students' vocational skills are developed in their personalized learning environment through collaborating with other students (Valtonen et al., 2012).

3 Method

3.1 *Context and Participants*

Twenty vocational teacher participants in four different subject areas (mechanical engineering, electronic engineering, international trade on e-commerce, and business English) were selected at two higher vocational colleges in Guangdong province, south China. These vocation-oriented subjects are closely related to local enterprises such as foreign trade companies and small- and medium-sized enterprises. Furthermore, higher vocational colleges in Guangdong province joined the scheme of industry-university collaboration to promote application-oriented teaching and students' vocational learning (Liu, 2016). Within the context of the demands of practice-oriented teaching and learning, the vocational learning environment includes flexible spaces for students' learning, adaptive instructions, and interactive classroom learning, which encourages learner engagement in the subject teaching. Each teacher participant has at least three years teaching experience. All teachers were observed three times during one teaching semester, and each observed class had around 25 to 30 students in one classroom. Therefore, 60 class observations (based on participants' agreement) were conducted using the International Comparative Analysis of Learning and Teaching (ICALT) instrument (Van de Grift, 2007, 2014) and the Comparative Analysis of Effective Teaching and Inspiring Teaching (CETIT) instrument (Ko et al., 2019).

3.2 *ICALT and CETIT Instruments*

Videotaped lesson observation was used to analyse students' interactive learning, teacher-student classroom interaction, and adaptive instructions in vocational learning environments that were embedded in the vocational pedagogy. The ICALT observation instrument has been applied to improving effective teaching behaviours and measuring teaching effectiveness and students' academic engagement in the Netherlands (Maulana & Helms-Lorenz, 2016; Maulana et al., 2017). The ICALT instrument was deemed appropriate to assess students' engagement and adaptive instructions within the vocational learning environment as it consists of six observable domains from the teacher's perspective: a safe and stimulating learning environment, efficient classroom management, clarity of instruction, activating teaching, the adaptation to students' learning needs, teaching learning strategies, and learner engagement from the student's perspective. Each domain comprises several indicators, and each indicator contains a number of items. For instance, the indicator of presenting and explaining the subject materials in the domain of clear and structured instructions includes items such as activating the prior knowledge of learners, giving staged instructions, posing questions which learners can understand, and summarising the subject material from time to time. Each item was rated on a 4-point

Likert scale (1 = mostly weak; 2 = more often weak than strong; 3 = more often strong than weak; 4 = mostly strong).

The CETIT observational instrument has similar features in terms of the domain of teaching behaviours when compared to the ICALT. The CETIT instrument covers 68 items in five aspects of inspiring teaching and employs a 5-point Likert scale in rating each item (1 = mostly weak; 2 = more often weak than strong; 3 = not observed (neutral); 4 = more often strong than weak; 5 = mostly strong). The CETIT observation instrument includes the features of teaching domains such as flexibility, collaboration, and innovative teaching that are more appropriate for vocational classrooms. For example, there are five items under the theme of classroom collaboration such as encouraging students to work together, giving students tasks to work in groups, students sharing their work in a task, making clear how students can help each other, and asking students to do demonstrations together.

3.3 Data Analysis

Two observation instruments (ICALT & CETIT) were employed to evaluate the quality of the lessons in terms of six aspects of effective teaching and five aspects of inspiring teaching, assuming these aspects occur independently. The mean scores of the two instruments were employed to rank the percentiles of the sixty lessons. Figure 9.1 summarizes the distribution of lessons' percentile rank based on their mean scores on each instrument, which were marked in red. The percentile rank

Percentile Rank of Lesson Mean on different scale
scatter plot

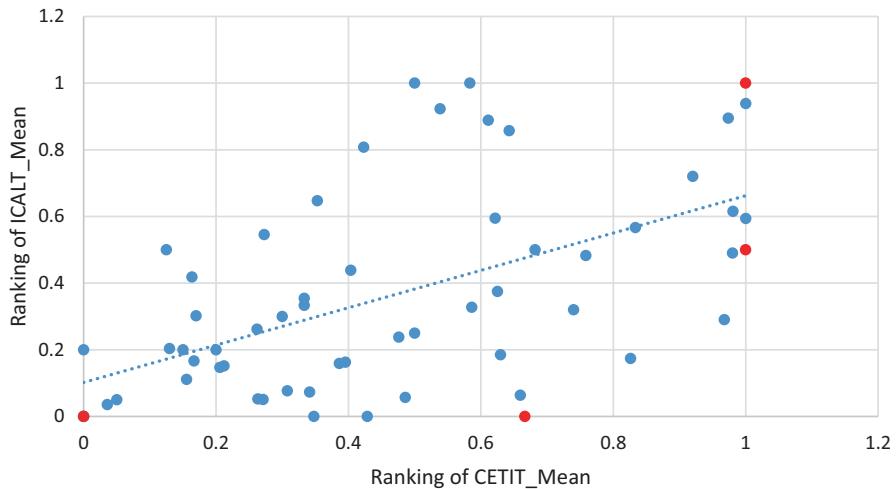


Fig. 9.1 The scatter plot of percentile rank of CETIT and ICALT mean scores

shows four contrastive cases: highly effective and highly inspiring, moderately effective and highly inspiring, moderately inspiring and highly ineffective, and highly ineffective and very uninspiring based on the mean scores of the distribution of lessons in Fig. 9.1. Four outlier lessons were selected to conduct an in-depth qualitative dialogue analysis (Hennessy et al., 2016; Hennessy et al., 2020; Vrikki et al., 2019b) to explore the teacher-student interactions in terms of the teaching effectiveness and the students' learning engagement.

A coding scheme for educational dialogue analysis (SEDA) developed by (Hennessy et al., 2016; Hennessy et al., 2020) consists of three hierarchical levels of analysis in a dialogic teaching and learning environment: communicative situations (CS) at a macro level, communicative events (CE) at a meso-level, and communicative acts (CA) at a micro-level. The SEDA coding scheme was used to analyse both the teacher's and the students' dynamic interactional process throughout a lesson according to these analytic procedures. Some studies have argued for the inclusion of various dialogue interactions that influence students' learning such as transactional distance dialogic interactions (Ekwunife-Orakwue & Teng, 2014), children's thinking and learning through dialogic approaches (Gillies, 2016, 2019; Maine & Hofmann, 2016; Rojas-Drummond et al., 2013), and students' linguistic development through dialogic teaching (Haneda, 2016; Haneda et al., 2017). In this study, we emphasised different forms of dialogic interactions in analysing four vocation-oriented lessons and the forms of dialogue are presented as follows:

Teacher-Led-Student ($T_L S$) interaction: dialogues between students and teachers – teacher-guided activities aiming towards increasing students' understanding.

Student-Led-Teacher ($S_L T$) interaction: dialogues between students and teachers – the teacher as a facilitator, student-led activities aiming towards increasing students' understanding.

Student-Student (SS) interaction: dialogues between students in group activities aiming towards increasing students' learning engagement.

Student-Content (SC) interaction: dialogues between students and course contents, that is, students' interaction with the technology or other materials used in the course or students' access to training platforms.

Four characteristic areas of vocation-oriented teaching and learning were summarised from the dimensions of the ICALT and CETIT instruments (structured and purposeful instructions, flexible and activating teaching, collaborative learning, and adaptive instructions) in relating to vocational students' learning engagement. In order to understand the selected dimensions of vocational teaching and learning and the general dynamics of the selected lesson(s), the CS was further segmented into a series of CE, i.e., each CS was segmented into different keyword descriptions as shown in Table 9.1.

Analysis of classroom dialogic interactions is an essential step in identifying a certain CE. CA, as a series of observable teacher-student and students' dialogic interactions were analysed using the coding scheme to code CA. In-depth analyses of videotaped lesson transcripts were carried out to describe vocational CS, CE, and CA under the forms of interactive dialogues ($T_L S$, $S_L T$, SS, SC). Table 9.2 below

Table 9.1 The related dimensions of ICALT and CETIT for dialogic analysis

Area code	Area name	Descriptions
S	Structured and purposeful instructions	<ol style="list-style-type: none"> 1. Explaining the learning objectives and purposes; 2. Explaining the subject material in a clear manner and giving feedback to learners; 3. Checking whether learners have understood the subject materials; 4. Giving a clear explanation of using didactic aids and carrying out assignments such as vocation or work-related activities; 5. Using materials and examples to illustrate the course content; 6. Encouraging students to make connections between what they learn with reference to their lives and to find different solutions for a problem.
C	Collaborative learning	<ol style="list-style-type: none"> 1. Encouraging students to work together; 2. Giving tasks/assignments to work in groups; 3. Guiding students to share their work in a task; 4. Making clear how students could help each other; 5. Asking students to do demonstrations together.
F	Flexible and active teaching	<ol style="list-style-type: none"> 1. Giving students some opportunities to choose their preferred classroom activities; 2. Allowing options for students in their own seat work and options for students in their homework; 3. Stimulating learners to think about solutions and ask questions 4. Stimulating learners to reflect and lets learners think aloud, 5. Giving interactive instructions and clearly specifying the lesson aims.
A	Adaptive instructions	<ol style="list-style-type: none"> 1. Adjusting his/her teaching pace or teaching methods to suit some students' needs, 2. Allowing for flexibility in the time learners get to complete assignments, 3. Giving weaker learners extra exercises/practice and 'pre- or post-instruction' 4. Using additional aids and means and giving additional instructions to small groups or individual learners 5. Adjusting the processing of subject matter to relevant inter-learner differences.
L	Learner engagement	<ol style="list-style-type: none"> 1. Paying attention as instructions are being given; 2. Participating actively in conversations and discussions and asking questions 3. Actively listening when instructions are being given; 4. Showing their interest by asking follow-up questions 5. Working independently 6. Taking the initiative themselves and using their time efficiently

Table 9.2 Excerpt from a three-minute dialogic analysis on automobile engineering about adding refrigerant

Agent	The teacher guides students on how to add the refrigerant to an automobile air conditioner	Teacher-led-Student interaction (T _L S)	Student-led-Teacher interaction (S _L T)	Student-Student interaction (SS)	Student-Content interaction (SC)
Teacher	We will add air-conditioning refrigerant today (students: um); firstly, as we said yesterday, this is a manifold pressure gauge (the teacher holds the pressure gauge on the right hand with a glove and uses his left finger to point at the manifold pressure gauge).	S1&S4			
Teacher	Then is it the red represents high pressure tube or low-pressure tube?	S3			
All students	That it is high pressure tube.				L2
Teacher	Then we need to open the lid of the high-pressure valve on the tube first (the student and the teacher open the lid together, the teacher releases the hand, while the student continues to twist the valve).	S4	C1		
Teacher & student	The teacher connects the blue high-pressure valve to the high-pressure place and screws it many times (the student asks if it is to ensure that the connection is all right; the teacher says: No, it is to ensure that it is closed; the student said: Yes, closed.)	S5	L4		L2
Teacher	When we hear sound, the vacuum indicator is already working, then at the same time to observe (while talking, the teacher points at the blue and red high and low table with hands) high- and low-pressure meter, the pointer of these two tables will go down, right?	S5&F3			
All students	Right.				L1

(continued)

Table 9.2 (continued)

Agent	The teacher guides students on how to add the refrigerant to an automobile air conditioner	Teacher-led-Student interaction (T _L S)	Student-led-Teacher interaction (S _L T)	Student-Student interaction (SS)	Student-Content interaction (SC)
Student A	Is it pumping refrigerant?		L4		
Student C	Another student points at the red valve which the teacher is screwing and says: you have screwed it in the wrong direction			L2	
Student B	Yes, it is the refrigerant.				L1
All students	Where is the refrigerant after vacuuming?		L2		
Teacher	In the air.	S2			
Teacher	That is the whole process of how to add the refrigerant, now it is your turn to do it again.	S3&C2			
Student A	No problem.				
Teacher	Now two students operate it together.	S3	C1		
Student B	Should it be turned off before the connections?			F3	L1
Teacher	Yes, it should be all turned off before any connections.	S2			
Student C	Ok, all done.			L6	
Teacher	After you have done all the connections, you have to double check it whether it is loose or not.	C4			
Student B	Yes, I have just tested it.				L2
Student A	Now it is adding not extracting.			L3	
Student D	Is that finished if we fill three bottles of refrigerant?			L4	
Teacher	Please open it first, when you hear some noises and it is starting to extract vacuum out, right? This side should make it larger.	F3			L3

shows a three-minute excerpt from a mechanical engineering lesson analysis which highlights the collaborative learning activities in the vocation-oriented training class.

4 Findings and Discussion

Dialogic teaching analysis of four videotaped lessons suggests that vocational teachers used informal and formal approaches to engage students in different aspects of classroom practice, such as small group teaching of vocational skills, vocation-oriented activities that connected to real-life situations, and students' collaborative learning on improving career competencies. The four lesson cases represent four different vocational majors that characterise students' collaborative learning and teachers' individualised or adaptive instructions. In practice, although the two selected lessons (from the subject area of international trade on e-commerce and electronic engineering respectively) show moderately inspiring and highly ineffective characteristics, and highly ineffective and very uninspiring characteristics based on the mean score distribution of the ICALT and CETIT instruments, there appears to be slight difference between the vocational lessons indicated regarding the collaborative learning and the vocation-oriented teaching and learning processes through dialogic analysis of vocational classroom interactions.

4.1 *Dialogic Teaching with Enhanced Learning Engagement*

Vocational students' engagement with purposeful instructions in small-group collaborative learning improved vocation-oriented teaching effectiveness through teacher-led classroom conversations and discussions. The automobile engineering lesson was set up 4 to 8 students in a group to operate the machine and the e-commerce lesson was formed of students supplied with installed e-commerce software for online interactive training. The extract detailing the teacher guided students working together to practice how to add refrigerant for automobile air conditioning shown in Table 9.2 suggests that the teacher-led classroom interactions emphasised clear and structured instructions in using materials relating to the course content in order to stimulate student-students learning engagement. The classes were featured as small group teaching and they were also designed as group teaching so that two teachers were guiding two groups of students and the other two groups were writing training reports or having their own practice within a group. The findings are informed by reviews of relevant literature (Gillies, 2019; Howe & Abedin, 2013; Lyle, 2008; Maine & Hofmann, 2016; Vrikki et al., 2019b) that dialogic small-group collaborative learning encourages students' involvement in vocational learning activities. Vocational students' collaborative learning in small groups promotes a stimulating learning environment that helps students improve

problem-solving skills (Hoek & Seegers, 2005). Meanwhile, Słowikowski et al. (2018) highlight that collaborative learning in online situations enhances students' vocational skills connecting with mechatronics education.

4.2 Adaptive Instructions on Vocation-Oriented Learning Activities

The pre-defined activities in the automobile engineering lesson allowed the teacher to adapt students' learning behaviors. For example, students who were falling behind in the pre-designed training activities were guided by using the other training machines. Moreover, in completing their class training activities, the teacher also used additional aids such as the flow chart board of operational procedures and a teaching assistant supporting them to handle the machines while the other grouped students were completing their after-training report assignments. According to Dillenbourg (2013), extrinsic activities are the main learning scenarios in classroom life and the core activities designed as adaptive with individualized instructions adapt the activities to students' learning. Based on the findings, adaptive vocational teaching featured individualized, structured, and purposeful instructions to adjust students' vocation-oriented learning activities. The adaptive instructions in this Chinese vocational learning environment meant that teachers could arrange task-based learning activities according to different subject requirements such as technological e-commerce platforms or training machines for engineering students. This suggests that differentiated instructions in vocation-oriented classrooms are directed towards engaging students' learning in various subject-based activities.

Although some studies have identified that differentiation in adjusting to learner differences is one of the more complex skills among teaching behaviours and student teachers and even experienced teachers spend a long time in developing this skill (Maulana & Helms-Lorenz, 2016; Van de Grift et al., 2014), other research finds that students' engagement in diversified vocational learning environments allows teachers to focus more on adapting to students' practical learning and vocational training (Zhao & Ko, 2020). On the practical level, however, adapting activities in the vocational classroom requires that teachers change the level of difficulty, such as adding or skipping some exercises whenever it is needed (Dillenbourg, 2013; Parsons, 2012). Therefore, adaptive instructions in the vocation-oriented training classroom attempt to integrate into learning environments while adjusting to both individualized and group learning activities. These findings are consistent with the view that a flexible classroom setting in the vocational learning environment promotes a stimulating learning climate and allows teachers to adjust pre-designed class activities in order to suit students' learning requirements (Dillenbourg, 2013; Dillenbourg & Tchounikine, 2007). This illustrated the flexibility of adaptive instructions within specialised vocational classroom activities and the focus on individual students' learning and instructions.

4.3 *Built-in Flexible Teaching with Enhanced Practical Understandings*

Flexibility was built into the vocational teaching arrangement in terms of the possibility of change while preparing class activities and the possibility of adjusting the teaching pace for some students to catch up with the average students. For example, the teacher guided students to work on the computer platforms themselves and walked around to help students in need and then gave them individual instructions in the e-commerce training class. It was evident that the flexible vocational learning environment allows teachers to modify interactive classroom activities (Dillenbourg, 1999, 2013) and collaborative lesson planning and teaching were characterized by the flexible nature of the learning environment and teaching and learning within open-plan settings (Deed et al., 2019). Furthermore, students in the electronic engineering class were flexibly arranged to perform classroom activities in order to promote interactions between students in completing their training projects. As is stated by Kuijpers et al. (2011) the flexible vocational learning environment fosters the development of students' career competencies, while students' vocational skills are emphasized in their personalized learning environment through collaborating with other students (Valtonen et al., 2012). The findings also supported the view that the development of practical learning achievement within individualized or fluid groupings was enhanced in the flexible learning environment (Deed et al., 2019).

5 Conclusion and Implications

This study illustrates students' learning engagement and teachers' adaptive instructions in a flexible vocational learning environment and group-based collaborative learning environments that support differentiated teaching practice and multiple class groupings of vocation-oriented activities. Furthermore, the study also reveals that flexibility in teaching stimulates vocational students' involvement in learning and promotes interactions between students' learning activities. In addition, the findings suggest that adaptations in vocational instructions may change based on students' engagement with learning activities as well as the flexibility of teaching scenarios. This supports previous research which demonstrates how teachers adapt their teaching practice to rely on the possibilities inherent in the flexible classroom environment and in their vocational instructions to engage their students and how this is more likely within specialized subjects (Deed et al., 2019; Dillenbourg et al., 2002; Dillenbourg, 2013; Zhao & Ko, 2020).

Although the research may be limited by the number of lessons analysed, the findings offer an in-depth understanding of vocational students' engagement in the collaborative learning environment and the flexibility provided by adaptations including structured and purposeful instructions in the vocation-oriented classroom. The study emphasises vocational students' learning patterns and teaching

adaptations in the specific context of Chinese higher vocational education, which suggests that vocational students' learning engagement and occupational competence development may be influenced by the collaborative learning environment and group-based adaptive instructions. Furthermore, the study contributes to the development of vocational learning theory and practice by enhancing our knowledge of student learning patterns, teaching practice, and the respective learning environments in the vocational education context. It also informs practitioners of the importance of vocational learning competence embedded in the delivery of vocational education curricula. Finally, vocation-oriented instructions involve teachers' workplace experiences while guiding students' training activities, which implies that vocational teachers' workplace learning experience may help improve collaborative learning activities and adaptive skill-based instructions in vocational classrooms.

References

- Alexander, R. J. (2004). *Towards dialogic teaching: Rethinking classroom talk* (1st ed.).
- Alexander, R. J. (2017). *Towards dialogic teaching: Rethinking classroom talk* (5th ed.). Dialogos.
- Alfassi, M. (2004). Effects of a learner-centred environment on the the academic competence and motivation of students at risk. *Learning Environments Research*, 7, 1–22.
- Arbaugh, J. B. (2000). How classroom environment and student engagement affect learning in Internet-based MBA courses. *Business Communication Quarterly*, 63(4), 9–26.
- Dane, J. (2016). The effective teaching and learning spatial framework. In W. Imms, B. Cleveland, & K. Fisher (Eds.), *Evaluating learning environments: Snapshots of emerging issues, methods and knowledge* (pp. 211–228). Sense.
- Darling-Hammond, L., & Bransford, J. (Eds.). (2007). *Preparing teachers for a changing world: What teachers should learn and be able to do*. Wiley.
- de Vries, S., Jansen, E. P., Helms-Lorenz, M., & van de Grift, W. J. (2015). Student teachers' participation in learning activities and effective teaching behaviours. *European Journal of Teacher Education*, 38(4), 460–483. <https://doi.org/10.1080/02619768.2015.1061990>
- Deed, C., Blake, D., Henriksen, J., Mooney, A., Prain, V., Tytler, R., et al. (2019). Teacher adaptation to flexible learning environments. *Learning Environments Research*. <https://doi.org/10.1007/s10984-019-09302-0>
- Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1–19). Elsevier.
- Dillenbourg, P. (2013). Design for classroom orchestration. *Computers and Education*, 69, 485–492. <https://doi.org/10.1016/j.compedu.2013.04.013>
- Dillenbourg, P., & Tchounikine, P. (2007). Flexibility in macro-scripts for computer supported collaborative learning. *Journal of Computer Assisted Learning*, 23(1), 1–13. <https://doi.org/10.1111/j.1365-2729.2007.00191.x>
- Dillenbourg, P., Schneider, D., Synteta, P., Dillenbourg, P., Schneider, D., Synteta, P., et al. (2002). Virtual learning environments. *Information & Communication Technologies in Education*, 3–18. <https://telearn.archives-ouvertes.fr/hal-00190701>
- Ekwunife-Orakwue, K. C., & Teng, T. L. (2014). The impact of transactional distance Dialogic interactions on student learning outcomes in online and blended environments. *Computers & Education*, 78, 414–427. <https://doi.org/10.1016/j.compedu.2014.06.011>
- Fraser, B. J. (2001). Twenty thousand hours: Editor's introduction. *Learning Environments Research*, 4(1), 1–5. <https://doi.org/10.1023/A:1011406709483>

- Gillies, R. M. (2016). Dialogic interactions in the cooperative classroom. *International Journal of Educational Research*, 76, 178–189. <https://doi.org/10.1016/j.ijer.2015.02.009>
- Gillies, R. M. (2019). Promoting academically productive student dialogue during collaborative learning. *International Journal of Educational Research*, 97, 200–209. <https://doi.org/10.1016/j.ijer.2017.07.014>
- Granger, E. M., Bevis, T. H., Saka, Y., Southerland, S. A., Sampson, V., & Tate, R. L. (2012). The efficacy of student-centered instruction in supporting science learning. *Science*, 338(6103), 105–108. <https://doi.org/10.1126/science.1223709>
- Haneda, M. (2016). Dialogic learning and teaching across diverse contexts: Promises and challenges. *Language and Education*, 31(1), 1–5. <https://doi.org/10.1080/09500782.2016.1230128>
- Haneda, M., Teemant, A., & Sherman, B. (2017). Instructional coaching through dialogic interaction: Helping a teacher to become agentive in her practice. *Language and Education*, 31(1), 46–64. <https://doi.org/10.1080/09500782.2016.1230127>
- Hardman, J. (2016). Opening-up classroom discourse to promote and enhance active, collaborative and cognitively-engaging student learning experiences. In C. Gorla, O. Speicher, & S. Stollhans (Eds.), *Innovative language teaching and learning at university: Enhancing participation and collaboration* (pp. 5–16). Research-publishing.net. <https://doi.org/10.14705/rpnet.2016.000400>
- Hardman, J. (2019). Developing and supporting implementation of a dialogic pedagogy in a primary schools in England. *Teaching and Teacher Education*, 86, 1–14. <https://doi.org/10.1016/j.tate.2019.102908>
- Hennessy, S., Rojas-Drummond, S., Higham, R., Márquez, A. M., Maine, F., Ríos, R. M., et al. (2016). Developing a coding scheme for analysing classroom dialogue across educational contexts. *Learning, Culture and Social Interaction*, 9, 16–44. <https://doi.org/10.1016/j.lcsi.2015.12.001>
- Hennessy, S., Howe, C., Mercer, N., & Vrikki, M. (2020). Coding classroom dialogue: Methodological considerations for researchers. *Learning, Culture and Social Interaction*, 25, 100404. <https://doi.org/10.1016/j.lcsi.2020.100404>
- Hoek, D. J., & Seegers, G. (2005). Effects of instruction on verbal interactions during collaborative problem solving. *Learning Environments Research*, 8, 19–39. <https://doi.org/10.1007/s10984-005-7949-9>
- Howe, C., & Abedin, M. (2013). Classroom dialogue: A systematic review across four decades of research. *Cambridge Journal of Education*, 43(3), 325–356. <https://doi.org/10.1080/0305764X.2013.786024>
- Kim, M. Y., & Wilkinson, I. A. (2019). What is dialogic teaching? Constructing, deconstructing, and reconstructing a pedagogy of classroom talk. *Learning, Culture and Social Interaction*, 21, 70–86. <https://doi.org/10.1016/j.lcsi.2019.02.003>
- Ko, J., Sammons, P., Maulana, R., Li, V., & Kyriakides, L. (2019, April 4–9). *Identifying inspiring versus effective teaching: How do they link and differ?* Paper presented at the American Educational Research Association Annual Meeting, Toronto, Canada.
- Kuijpers, M., Meijers, F., & Gundy, C. (2011). The relationship between learning environment and career competencies of students in vocational education. *Journal of Vocational Behavior*, 78(1), 21–30. <https://doi.org/10.1016/j.jvb.2010.05.005>
- Kumpulainen, K., & Lipponen, L. (2010). Productive interaction as agentic participation in dialogic enquiry. Educational dialogues: Understanding and promoting productive interaction. In K. Littleton & C. Howe (Eds.), *Educational dialogues. Understanding and promoting productive interaction* (pp. 48–63). Routledge. <https://doi.org/10.4324/9780203863510>
- Liu, Y. (2016). Province/ministry-coordinated industry-university-institute cooperation and university development: Based on the experiences of Guangdong Province. *Chinese Education and Society*, 49(3), 139–151. <https://doi.org/10.1080/10611932.2016.1218252>
- Lyle, S. (2008). Dialogic teaching: Discussing theoretical contexts and reviewing evidence from classroom practice. *Language and Education*, 22(3), 222–240. <https://doi.org/10.1080/09500780802152499>

- Maine, F., & Hofmann, R. (2016). Talking for meaning: The dialogic engagement of teachers and children in a small group reading context. *International Journal of Educational Research*, 75, 45–56. <https://doi.org/10.1016/j.ijer.2015.10.007>
- Maulana, R., & Helms-Lorenz, M. (2016). Observations and student perceptions of the quality of preservice teachers' teaching behaviour: Construct representation and predictive quality. *Learning Environments Research*, 19(3), 335–357. <https://doi.org/10.1007/s10984-016-9215-8>
- Maulana, R., Helms-Lorenz, M., & Van de Grift, W. (2017). Validating a model of effective teaching behaviour of preservice teachers. *Teachers and Teaching*, 23(4), 471–493. <https://doi.org/10.1080/13540602.2016.1211102>
- Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A socio-cultural approach*. Routledge.
- Nystrand, M., Wu, L. L., Gamoran, A., Zeiser, S., & Long, D. A. (2003). Questions in time: Investigating the structure and dynamics of unfolding classroom discourse. *Discourse Processes*, 35(2), 135–198. https://doi.org/10.1207/S15326950DP3502_3
- Parsons, S. A. (2012). Adaptive teaching in literacy instruction: Case studies of two teachers. *Journal of Literacy Research*, 44(2), 149–170. <https://doi.org/10.1177/1086296X12440261>
- Rojas-Drummond, S., Torreblanca, O., Pedraza, H., Vélez, M., & Guzmán, K. (2013). 'Dialogic scaffolding': Enhancing learning and understanding in collaborative contexts. *Learning, Culture and Social Interaction*, 2(1), 11–21. <https://doi.org/10.1016/j.lcsi.2012.12.003>
- Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77(4), 454–499. <https://doi.org/10.3102/0034654307310317>
- Słowikowski, M., Pilat, Z., Smater, M., & Zieliński, J. (2018, October). Collaborative learning environment in vocational education. In *AIP conference proceedings* (Vol. 2029, No. 1, p. 020070). AIP Publishing LLC.
- Teo, P. (2016). Exploring the dialogic space in teaching: A study of teacher talk in the pre-university classroom in Singapore. *Teaching and Teacher Education*, 56, 47–60. <https://doi.org/10.1016/j.tate.2016.01.019>
- Valtonen, T., Hacklin, S., Dillon, P., Vesisenaho, M., Kukkonen, J., & Hietanen, A. (2012). Perspectives on personal learning environments held by vocational students. *Computers & Education*, 58(2), 732–739. <https://doi.org/10.1016/j.compedu.2011.09.025>
- Van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research*, 49(2), 127–152. <https://doi.org/10.1080/00131880701369651>
- Van de Grift, W. J. (2014). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement*, 25(3), 295–311. <https://doi.org/10.1080/09243453.2013.794845>
- Van de Grift, W., Helms-Lorenz, M., & Maulana, R. (2014). Teaching skills of student teachers: Calibration of an evaluation instrument and its value in predicting student academic engagement. *Studies in Educational Evaluation*, 43, 150–159. <https://doi.org/10.1016/j.stueduc.2014.09.003>
- Vriikki, M., Wheatley, L., Howe, C., Hennessy, S., & Mercer, N. (2019a). Dialogic practices in primary school classrooms. *Language and Education*, 33(1), 85–100. <https://doi.org/10.1080/09500782.2018.1509988>
- Vriikki, M., Kershner, R., Calcagni, E., Hennessy, S., Lee, L., Hernández, F., et al. (2019b). The teacher scheme for educational dialogue analysis (T-SEDA): Developing a research-based observation tool for supporting teacher inquiry into pupils' participation in classroom dialogue. *International Journal of Research & Method in Education*, 42(2), 185–203. <https://doi.org/10.1080/1743727X.2018.1467890>
- Zhao, Y., & Ko, J. (2020). How do teaching quality and pedagogical practice enhance vocational student engagement? A mixed-method classroom observation approach. *International Journal of Educational Management*, 34(6), 987–1000. <https://www.emerald.com/insight/0951-354X.htm>

Yanmin Zhao is currently a Lecturer and the Doctor of Education Programme Coordinator at the Graduate School of the Education University of Hong Kong. She is interested in the field of classroom research and the changing learning environment of vocational students. Before her doctoral study, Yanmin was an EFL teacher for about four years and had an experience in student discipline, counselling and guidance, curriculum development, and teacher professional development. Her current research focuses on the professional learning of teachers, pedagogical practice and workplace learning in applied degree education.

Marc Kleinknecht is Chair Professor of Teacher Education and School Development at the Leuphana University of Lüneburg, Germany; Degree as Teacher (2000) and in Education (2005); PhD. in Educational Science (2010) and Venia Legendi for Educational Science from Technical University of Munich (2016). Research interests: Teaching Quality, Video-based Teacher Learning, Practice-based Teacher Education. In current studies, he is investigating the impact of classroom-video-based feedback of peers and experts on teachers' professional vision and teaching practices.

Dr. James Ko is an Associate Professor at the Department of Policy Leadership and Co-Director of the Joseph Lau Luen Hung Charitable Trust Asia Pacific Centre for Leadership and Change at the Education University of Hong Kong. Before his doctoral study, James was an EFL teacher for about 20 years and led two functional teams in a secondary school for 10 years. He is a recurrent grantee of the RGC and UGC grants and the principal investigator of 23 projects, collaborating with local academics and overseas researchers on 40 projects. He has supervised 14 doctoral students with 8 completed.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 10

Teaching Quality in Indonesia: What Needs to Be Improved?



Yulia Irnidayanti and Nurul Fadhilah

Abstract Based on international testing results (e.g., PISA, 2015; TIMSS, 2015), the performance of Indonesian students remains poor. The low quality of education in Indonesia is determined by many factors, including the teacher's quality. Teachers have a very strategic role in the learning process. Effective teaching behavior is used as an indicator of teaching quality and is the main target of this study, which is needed to improve the teaching quality of teachers in Indonesia. Research on effective, evidence-based, teaching behavior has identified six domains of effective teaching behavior, which are relevant to the Indonesian context. In this chapter, we will describe Indonesian secondary school teachers' teaching behavior based on trained observers' and students' reports. The ICALT and My Teacher Questionnaire were used to gather data across 13 provinces in Indonesia, covering about 375 teachers and 6410 students. The quality level of effective teaching behavior was examined, and similarities and differences between observers and student reports were discussed. This study result shows the profile of teacher teaching quality in Indonesia that can be used as a basis for policy making related to improving teaching and professional development of teachers in Indonesia.

Keywords Teaching quality · Indonesia · Teacher · Observer · Student perception · Differentiated instruction

Y. Irnidayanti (✉)

Department of Biology and Biology Education, Faculty of Mathematics and Science,
State University of Jakarta, Jakarta, Indonesia

N. Fadhilah

Department of Biostatistics and Population Studies, Faculty of Public Health,
Universitas Indonesia, Depok City, Indonesia

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_10

225

1 Introduction

Two large-scale comparative assessments organized by the Organization for Economic Cooperation and Development (OECD) and the International Association for the Evaluation of Educational Achievement (IEA) have provided useful insights into trends in educational performance around the world (Martin et al., 2016; Mullis et al., 2016; OECD, 2015). Trends in education outcomes show that Indonesia consistently ranks among the lowest performers. One of the many factors that play an important role in the low quality of education in Indonesia is the quality of teachers. Teacher quality is influenced by qualifications such as teacher education level, teaching experience, participation in professional development activities, and self-efficacy (Goe, 2007). Teacher quality has been shown to be critical to student achievement (Baumert et al., 2010; Blömeke & Delaney, 2014) and is strongly linked to teaching quality. All these variables are the most important factors for student learning at the classroom level (Kyriakides et al., 2009).

Teacher quality is a construct, which reflects the characteristics of teacher teaching practices that are positively related to student learning outcomes, both cognitive and affective (Maulana & Helms-Lorenz, 2016). The quality of effective teaching is reflected in the teaching behaviour of teachers in the classroom.

In the 1980s, due largely to changes in economic, social, and educational developments around the world, teachers began to be expected to learn during their careers (Beijaard et al., 2007) and teachers were expected to become “adaptive experts” in the learning process (Vermunt & Verloop, 1999; Wei et al., 2009). Teacher learning throughout the career is related to improving teaching practices. In response to these insights, improvement in teaching quality via teaching practices has been included on the professional development agenda for teachers in many countries.

In Indonesia, teacher professional development programme has been carried out since 2005 through/being the PPG (Teacher Professional Education) program, PLPG (Teacher Professional Education and Training) and UKG (Teacher Competency Test) (Kemendikbud, 2016). Nevertheless, Indonesia is remains lowest in the ranking of Asia as well as Europe. The recent research about effective teaching behaviour across six countries (Netherlands, Spain, Turkey, South Africa, South Korea, and Indonesia) based on student perception’s shows that perceived teaching behaviour was the highest in South Korea and the lowest in Indonesia (André et al., 2020). Another recent research, related to teaching behaviour across various national contexts based on the observer’s perception in each country, including the Netherlands, South Korea, South Africa, Indonesia, Hong Kong-China, and Pakistan, indicates South Korea always the highest quality of teaching behaviour, while Indonesia ranked the lowest (Maulana et al., 2020). Hence, differences in the quality of teaching practices may partly explain differences in countries’ average educational outcomes. Other issues, including teacher motivation, teacher selection, and initial teacher training programs have been put forward contributing factors to the low quality of education in Indonesia (De Ree, 2016; Fasih et al., 2018).

Based on Law no. 14 of 2005 the basic competencies that must be possessed by a teacher in Indonesia are pedagogic, personality, professional, and social competencies. Pedagogic competence includes the ability to plan, implement, manage, and evaluate the learning process, as well as being able to understand and actualize students with various potentials. These basic competencies are not only a requirement to become a teacher but must be implemented in learning activities in the classroom. Effective teaching behaviour as an indicator of teaching quality is the main target of this research. Research on evidence-based effective teaching behaviour has identified six domains of effective teaching behaviour (Van de Grift, 2007) relevant to the Indonesian context. This research conducted is relevant to the needs of the Indonesian government to measure teaching effectiveness in Indonesia. In this study, six domains of teaching quality will be observed, both based on the perception of trained observers using the ICALT observation instrument (van de Grift et al., 2014) and the perception of Indonesian students using the My Teacher Questionnaire (Maulana & Helms-Lorenz, 2016).

2 Theoretical Framework

2.1 Teaching Quality

Teachers play a very strategic role in increasing students' situational interest in active learning classroom (Rotgans & Schmidt, 2011), as well as participating in the curriculum planning process (Ben-Peretz, 1980). Therefore, the quality of education is highly dependent on the quality of the teacher, where the teacher is seen as a central figure in improving student academic performance to the highest level. Improving the quality of teachers is a work plan from the Indonesian Ministry of Education and Culture (2005–2025). Findings from the research indicated that teacher quality is associated with students' performance. Good teachers do not only display their competence in the subject area but also support their students in terms of displaying friendliness, optimism and creating a conducive learning environment (Hamid et al., 2012). Good quality teachers demonstrate effectiveness in teaching and have an impact on student achievement (Rice, 2003).

Evaluation of teacher quality can be analyzed using three approaches: input, process, and output. Inputs are what a teacher brings to his or her position, such as measured as teacher background, beliefs, expectations, experience, pedagogical and content knowledge, certification and licensure, and educational attainment. In the literature known as “teacher quality”. Processes refer to the interaction that occurs in a classroom between teachers and students. Outputs represent the results of the activity process in the classroom, such as the impact on student achievement, graduation rates, student behavior, engagement, attitudes, and social-emotional well-being. Goe et al. (2008) showed that outputs can be referred to as “teacher effectiveness,” as used in the research literature is often limited to the meaningful

impact on student achievement specifically. The five points of the effective teacher are defined as follows: (1) effective teachers have high expectations for all students and help students learn, (2) effective teachers contribute to positive academic, attitudinal, and social outcomes for students, (3) effective teachers use diverse resources to plan and structure engaging learning opportunities; monitor student progress using formative assessment, adapting instruction as needed; and evaluate learning using multiple sources of evidence, (4) effective teachers contribute to the development of classrooms and schools, (5) effective teachers collaborate with other teachers, administrators, parents, and education professionals to ensure student success.

Goldhaber (2015) stated that empirical research has shown that teacher quality is the largest in-schools factor that contributes to student achievement but the visible characteristics such as education level and certification status did not include. Variations in effective teaching behavior are usually categorized into and/or summarized by five to seven factors or broader domains (Muijs et al., 2014). The teaching behaviors used in this research are grouped into six domains, namely: safe and stimulating learning climate, efficient classroom management, clear and structured instructions, Intensive and activating teaching, teaching-learning strategies, and adaptation of teaching/differentiation (Van de Grift, 2007).

Examples of safe and stimulating learning climate practices are emphasizing on things such as creating a safe and relaxed and conducive learning atmosphere, stimulating students' self-confidence, stimulating motivation in learning, appreciating student work, always fostering solidarity among students, encouraging students to work in groups, creating a safe learning atmosphere, respecting students, and teachers. These aspects are also incorporated in the ICALT observation instrument and applicable to the learning climate of Indonesian schools (Maulana et al., 2015a).

Efficient classroom management is an important factor in supporting the creation of a safe and stimulating learning. It is an indispensable aspect of teaching quality (Harrell et al., 2004). Efficient in managing classrooms so as not to waste time studying. For this example of teaching practice, the teacher must begin and end the lesson on time, pay attention to the time transition, minimize wasting time during learning, such as not discussing things outside the context of the lesson, using time as efficiently as possible. This needs to be considered because lesson time is not always supported for learning activities but is often used for non-curricular activities, organizational matters or dealing with disciplinary problems (Kunter et al., 2007). Classroom organization and learning plans to use effective time are especially important where students are exposed to maximum learning opportunities (Wang et al., 1993).

Clear and structured instructions emphasize the concept of learning structure is clear and effective. Students are expected to be able to process information and to perform adequately (Gagne & Briggs, 1974). Learning instructions use clear and structured sentences, the subject matter is abstract, and complex should be made real and simplified. At the beginning of a lesson, the teacher must ensure that all students know what is expected of them at the end of the lesson by clearly stating

the lesson outcomes (Todd & Mason, 2005). Therefore, the subject matter should be clear and understandable; students should receive regular feedback to establish their progress; all students should be actively engaged in the lesson; the teacher must allow students to think, the teacher should explain in a well-structured manner and use didactic while explaining new concepts (Maulana et al., 2015b). Clear instruction can also be supported by how the teachers implement the curriculum, apply content to students' everyday life situations, and use language that is understandable to them (Vandeyar & Killen, 2007).

Intensive and activating teaching emphasizes the concept of continuous and interactive learning, using concepts and skills relevant to students' everyday lives (Downer et al., 2007). Teachers must actively ask, analyze and reason; give feedback in a way that stimulates student's efforts to learn. For the domain of intensive and activating teaching to be achieved, teachers must create and develop frameworks that can explore the potential that exists in students and provide motivation to build confidence in weak students, provide interactive instruction where they can collaboratively work with others in finding solutions to problems (Van de Grift, 2007).

Adaptation of teaching (differentiation) is described as learning following how to process between students. Heterogeneity of students must be facilitated during the learning process in classrooms. Therefore, a differentiated instruction framework is needed, such as providing free time to help weak students during learning, assigning different tasks between students, providing diverse activities, maximizing student potential in a variety of ways that are adapted to students. Differentiated instruction requires teachers to be mindful of the diverse characteristics of students in their classrooms. It refers to teaching behaviors including the adjustment of instruction and student processing to individual students according to differences in their learning profiles, learning needs and motivation (Pearson & Fielding, 1991). Differentiation instruction is very flexible, organized, and proactive. It can accommodate a variety of student learning preferences in achieving their full potential (Lawrence-Brown, 2004).

Domain teaching-learning strategy is needed to achieve student academic success. Cognitive and metacognitive strategies have a positive effect on student learning (Montague & Dietz, 2009). Cognitive strategies aim to help students achieve certain goals while metacognitive strategies precede cognitive activities to ensure that goals have been achieved (Roberts & Erdos, 1993). The cognitive approach is very efficient, where students are guided so that they are motivated to carry out activities independently (Pressley et al., 1990). These strategies can help students to connect new concepts with what they already know, besides helping them carry out higher-level procedures. Teachers who provide their students with learning strategies have a significant impact on their learning performance (Houtveen & van de Grift, 2007). Empirical confirmation of these six domains of teaching has been provided by Maulana et al., (2017a) and Irnidayanti and Fadhilah (2018).

2.2 The Profile of the Indonesian Teacher: Context for the Current Study

Recent research also supports that the quality of teacher in Indonesia is still low compared to other countries. Teaching behaviour based on the perception by students in Indonesia lower than the Netherlands, Spain, Turkey, South Africa, and South Korea (André et al., 2020; Maulana et al., 2020). Most of the teachers observed in this study were certified teachers, whose teaching quality was still low. These certified teachers do not apply their skills and competencies in the classroom (De Ree et al., 2018). Based on our research, teaching behavior is correlated with students' academic engagement. Teachers have not been optimal in involving students in the learning process. This can be seen from the results of our study which showed a moderate level of student involvement. Most teachers in Indonesia use a teacher-centered approach in the learning process. In the Asian context, particularly in Indonesia, pervasive cultural values are linked to power distance, which allows growth among people in hierarchies. This situation is reflected in the classroom where the teacher is the center (CIA, 2017).

2.3 Observer Perceptions of Teaching Quality

Teacher quality can be observed in their teaching behavior in the classroom. In general, there are three common tools for measuring teaching behavior: classroom observations, student surveys, and teacher surveys. Class observations can only be conducted by trained observers, where they assess what is happening in the classroom and the assessment is not influenced by students and teachers (Lawrenz et al., 2003). Classroom observations are viewed as the most objective in teaching practice (Worthen et al., 1997) and more often used than student surveys and teacher surveys (Goe et al., 2008).

The weakness of classroom observations is that the presence of an observer can influence teacher behavior in teaching practice (de Jong & Westerhof, 2001), which allows measurement of teaching behavior to be less accurate. In addition, classroom observations are very demanding and time-consuming because observers must be trained intensively and observations are made several times to get an objective and accurate measure of teaching behavior (Hill et al., 2012; van der Lans et al., 2015).

2.4 Student Perceptions of Teaching Quality

Students' perceptions are views or interpretations of students regarding interactions in learning activities in the classroom. Perceptions between students are different on the teaching behavior of teachers in the classroom. Assessment of teacher teaching

behavior based on students' perceptions contributes to the understanding of the quality of teaching in the classroom and is an important part compared to the assessment by outside observers. Student experiences in the classrooms conducted from time to time during learning involve their academic activities (den Brok et al., 2004). The evidence shows that most students' perceptions of teaching behavior are better predictor of learning outcomes compared that of a trained observer (De Jong & Westerhof, 2001; Seidel & Shavelson, 2007). Student and teacher surveys are known to be cost-effective and less demanding, and less time-consuming for measuring teaching behavior (Goe et al., 2008).

Students' perceptions at the classroom level are more valid and can predict and evaluate teaching behavior than external observers (Kyriakides, 2005; Goe et al., 2008). Student perceptions and teacher perceptions are related to the construct of teaching behavior (Kunter et al., 2008). There are some weaknesses related to student perceptions of teaching practices in the classroom. Students' perception can be influenced by various factors including their interpersonal closeness with their teachers, interest in the subject taught by their teachers, expectations about their grades, and student age (Peterson et al., 2000; Richardson, 2005; Benton & Cashin, 2012). Although students' perceptions have some weaknesses, the student evaluation of teaching has been one of the most widely used indicators of teacher effectiveness and educational quality (Scherer et al., 2016). De Jong and Westerhof (2001) and Seidel and Shavelson (2007) indicate that student perceptions are more predictive of student learning outcomes than external observations and teacher perceptions. Student perceptions should be considered although there are doubts about it regarding the objective assessment (Van de Grift, 2007). Student's perceptions could be useful when the focus of the assessment is the teaching strategies used in the classroom, the content subject, or the effectiveness of their teaching (Martínez-Rizo, 2012).

3 Aims of the Present Study

Research about the importance of teaching quality in developing countries, such as in Indonesia, is still very limited and scarce. Therefore, this research is needed to provide an overview of the quality of teaching and as evidence to find out and measure the quality of education in Indonesia. To guide the study, the following research questions were formulated:

1. How is the general profile of teaching quality of Indonesian perceived by their students and trained observers in terms of effective teaching behavior?
2. Can the general profile of teaching quality in Indonesia contribute to policy recommendations for the Indonesian educational system?
3. What needs to be improved in the teaching quality in Indonesia?

4 Methods

4.1 *Sample and Procedure*

The Indonesian sample used to measure the actual teaching behavior of teachers in the classroom consists of 375 teachers, who teach in 24 secondary schools in 13 provinces. The teacher sample came from varied socioeconomic backgrounds and different cultures. The sample consisted of 89.7% of teachers from public schools and the remaining teachers from vocational schools and private schools. The demographic distribution of the sample is as follows: 27.5% of schools were outside Java, 38.7% were Science related subjects, 41.6% were male teachers, 79.5% were experienced teachers, and 85.6% had large class sizes, 60.1% were female students. All schools are in various provinces: Pidie and Bireun (NAD), Lampung, Makassar (South Sulawesi), Bontang (Borneo), Tangerang (Banten), Bandung, Bekasi, Depok and Bogor (west java), Pekalongan and Wonosobo (central Java), Gresik (east Java), and Jakarta. A total of 6410 students was used to measure pupil's perception of teacher's teaching behavior. The percentage of missing cases is very low (< 0.5%), which indicates a very high response rate.

This study used direct classroom observation methods by trained observers and student surveys to assess teacher teaching behavior in natural environments using a validated instrument of ICALT observation and My Teacher's questionnaire (Maulana & Helms-Lorenz, 2016). Typical lessons from teachers are visited and observed by trained observers after an agreement is reached between researchers, schools, and teachers. The teachers and schools participated in this research voluntarily.

Schools were recruited to participate in the survey voluntarily. An agreement between the researcher-the school was made before conducting a survey in these schools. Letters were sent to the principals of the schools to participate in this research. Upon official agreement to participate, observations were conducted based on appointments during the school year. The survey involved 10 trained observers who traveled and observed the school mentioned above. The filling out of the questionnaire was conducted by trained observers to assess the actual learning process in the classroom, while the student survey was conducted after learning was completed to assess the teaching practices of their teachers. The time needed for students to fill out the questionnaire takes about 30 min to complete. After filling out the questionnaire was completed and was collected by the observer.

4.2 *Measuring Teaching Behaviour*

The validated Indonesian version of the International Comparative Analysis of Learning and Teaching (ICALT) observation instrument was used in this research to measure actual teachers' teaching behavior based on the observer (Maulana et al.,

2017b; Van de Grift et al., 2014). The reliability of ICALT observation instrument measured with Cronbach's alpha values ranging from 0.71–0.86, Scale reliability learning climate (0.710), Classroom management (0.77), Clarity of instruction (0.84), activating learning (0.81), adaptive instruction (0.81), teaching-learning strategies (0.86). ICALT observation instrument consists of 32 items, using four ordinal response categories (1 = 'mostly weak' to 4 = 'mostly strong').

We used the My Teacher Questionnaire (Maulana & Helms-Lorenz, 2016) based on the teaching behavior model of Van de Grift (2007) and Van de Grift et al. (2014). The instrument has proved to accurately measure teachers' teaching behavior based on student perceptions and the validated Indonesia version was used in this research. The total items of instrument MTQ is 41 items and the reliability of the ICALT observation instrument measured with Cronbach's alpha values ranging from 0.70–0.76. The instruments were translated to Indonesia and back translated for use in Indonesia based on the guidelines provided by Hambleton et al. (2004).

4.3 Data Analysis

Preliminary assumption testing was conducted to check for normality, homogeneity of variance, validity, and reliability of the instrument. To answer the first research question, descriptive analyses were calculated to determine the mean scores of teaching behavior, to get the general profile of teaching quality. To answer the second question, we analyzed descriptively the profile of teaching quality in Indonesia and other countries. We suggest on how to improve teaching quality in Indonesia based on related reference.

5 Results

5.1 General Profile of Teachers' Teaching Quality of Indonesian Perceived by Trained Observers and Their Students

Based on the ICALT observation instrument results, the level of effective teaching behavior in Indonesia is moderate/sufficient except for the differentiation instruction domain that is low/insufficient. The mean score of 6 domain teaching behavior based on the ICALT questionnaire are Safe and stimulating learning climate (2.88 ± 0.49), efficient classroom management (2.59 ± 0.65), Clear and structured instructions (2.45 ± 0.69), Intensive and activating teaching (2.31 ± 0.58), differentiated instruction (1.74 ± 0.68) and teaching-learning strategies (2.04 ± 0.62). Meanwhile, based on the student's My Teacher questionnaire, all six domains of the

level of effective teaching behavior in Indonesia is moderate/sufficient with the mean (\bar{x}) score ranging from 2.8 to 3.0.

The profile of teacher behavior in Indonesia based on observer perceptions shows that the adaptation of teaching (differentiation) is insufficient, while the remaining five (Safe and stimulating learning climate, Efficient classroom management, Clear and structured instructions, Intensive and activating teaching, differentiated instruction, Teaching-learning strategies), were rated as sufficient (Fig. 10.1). The quality of teachers plays an important role in determining the educational competitiveness of a country, especially in the era of globalization. Indonesia has recognized the importance of improving the quality of education, especially the quality of teachers.

In the Indonesian context, the lowest score of the six domains of teaching behavior is teaching adaptation (differentiation), with a score of 1.74 out of 4. Teaching and learning strategies are the second-lowest score on the profile of teaching behavior in Indonesia. Teaching and learning strategies are closely related to teaching adaptation (differentiated instruction). Learning in Indonesia is mostly a teacher-centered approach, where teachers usually provide the same teaching for all students. This approach is not suitable in the context of differentiation, which the teacher must be able to adapt to the needs of students in the classroom (World Bank, 2016; Tomlinson, 1999). The teacher makes distinctions in the classroom by

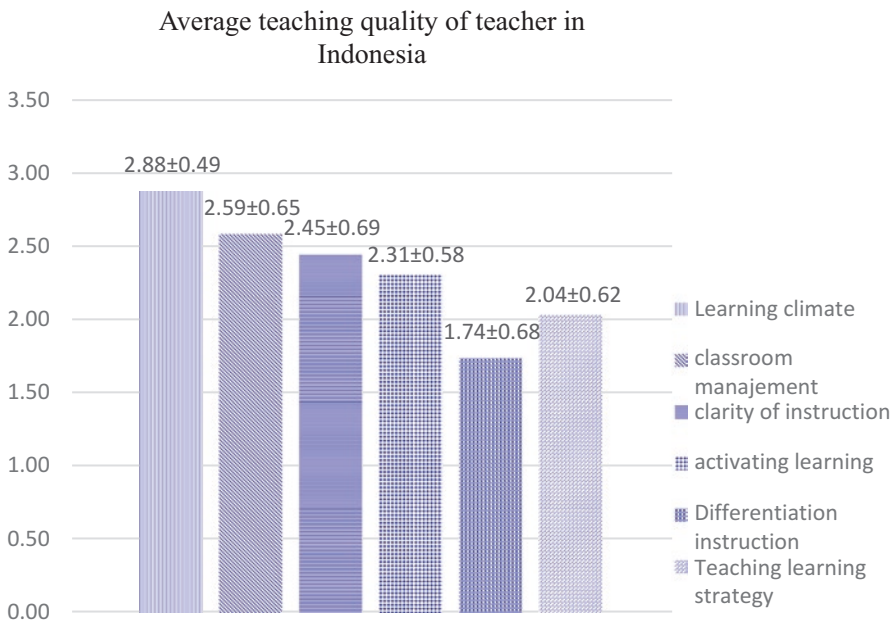


Fig. 10.1 The general profile of teacher’s teaching quality in Indonesia seen by Indonesian observer perception. Learning climate: Sufficient/Moderate, Classroom management: Sufficient/Moderate, Clarity of instruction: Sufficient/Moderate, activating learning: Sufficient/Moderate, differentiated instruction: Insufficient/poor, Teaching learning strategies: Sufficient/Moderate. Metric criteria: 1–1.99 = Insufficient/poor, 2.00–2.99 = Sufficient/moderate, 3.00–4.00 = Good/high

making discriminatory instructions. An example of a complex approach to teaching and learning is differentiated instruction. A model of teaching-learning strategy approach that serves various learning profiles is referred to as differentiation (Tomlinson, 2005; Subban, 2006).

The profile of teacher teaching quality in Indonesia based on student perceptions can be seen in Fig. 10.2. Results of descriptive analyses show that mean scores and the corresponding standard deviations for all domains are Safe and stimulating learning climate ($M = 2.93, SD = 0.45$), Efficient classroom management ($M = 3.05, SD = 0.39$), Clear and structured instructions ($M = 2.97, SD = 0.43$), Intensive and activating teaching ($M = 2.95, SD = 0.41$), differentiated instruction ($M = 2.88, SD = 0.45$), and Teaching learning strategies ($M = 2.83, SD = 0.43$). On average, teachers' classroom management was perceived as good, while the remaining five teaching behavior domains were rated as sufficient.

There are different perceptions regarding the general profile of teacher teaching quality in Indonesia between students and observers. The efficient classroom management is good based on students' perceptions, while the category is sufficient for efficient classroom management based on observer perceptions. Differences about perception also exist in the differentiated instruction. Based on the student's perception that the differentiated instruction is sufficient but based on the perception of the observer shows learning differentiation is insufficient.

Several factors contribute to the differences between observers and student's perception of the teacher's teaching behavior. The central participants in the classroom are the teacher and the student. The teacher arranges and creates the learning situation, which the student must accept. However, the success and effectiveness of the

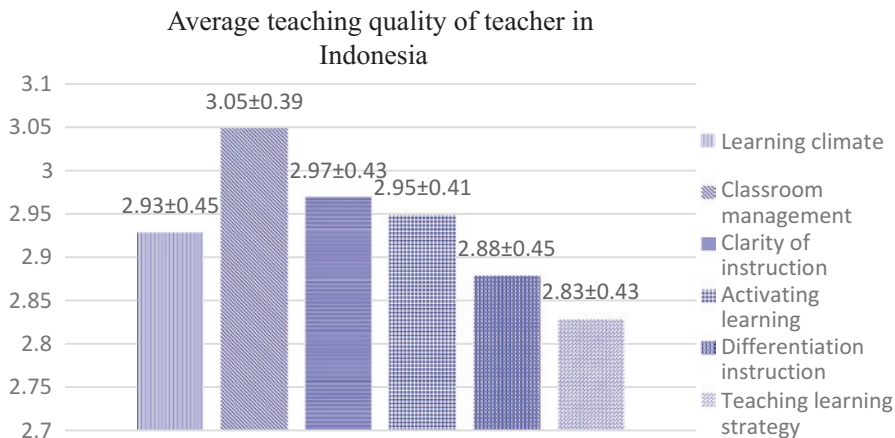


Fig. 10.2 The general profile of teacher's teaching quality in Indonesia seen by Indonesian student perception. Learning climate: sufficient/moderate, classroom management: good/high, clarity of instruction: sufficient/moderate, activating learning: sufficient/moderate, adaptive instruction: sufficient/moderate, teaching learning strategies: sufficient/moderate. Metric criteria: 1–1.99 = Insufficient/poor, 2.00–2.99 = Sufficient/moderate, 3.00–4.00 = Good/high

instruction depend on both parties (Fend, 2002). The student has a different role and has a different perspective with their teacher in the classroom. In this perspective, both teacher and student provide insight into what happens in the classroom (den Brok et al., 2006). The student has more time to observe ongoing classroom processes. Therefore, they have a broad base of experiences over many class hours with a variety of teachers.

Their judgements of their teacher are more consistent than external observers and teachers' judgement (den Brok et al., 2006). Students are an "excellent source" of information about classroom processes (Montuoro & Lewis, 2014). Sometimes, student's perceptions about their teachers reflect their subject knowledge comprehension because perception of student is individual perception and students don't have methodological-didactic knowledge (Wagner et al., 2016). Therefore, judgements on teaching behavior by external observers are better than the student's perception (Scherzinger & Wettstein, 2019). The external observers make comprehensible judgements and guided by rules. Because they are not involved in the interaction in the classroom, so their judgment is more objective (Praetorius et al., 2012).

6 Can the General Profile of Teaching Quality in Indonesia Contribute to Policy Recommendations for the Indonesian Educational System?

The profile of teaching quality in Indonesia is mostly sufficient except in differentiated instruction. However, in general the profile of teaching quality is lower than other country, such as Spain, Turkey, Netherland, South Korea, and South Africa (André et al., 2020), Hongkong -China, Pakistan (Maulana et al., 2020). There are several factors that cause the low teaching quality in Indonesia. Teachers' content knowledge is particularly important in determining student performance, while many teachers in Indonesia have very low content knowledge. Teachers with formal qualifications, such as a bachelor's degree, only have slightly better quality. The result of national civil service teachers' examination also shows the low quality of teacher candidate in Indonesia (World Bank, 2016). About 65% of the total of 2.7 million teachers in Indonesian, do not meet the requirements posed for professional teachers. The weakness of the national teacher training system results in the low quality of teacher candidates. This condition also influences the motivation of the lower ability teachers. They are reluctant to upgrade their skills and qualification (Jalal et al., 2009).

Another reason is the ineffective allocation of the education budget. The allocation of Indonesian education funds is only used for teacher allowances and unfortunately, the large allocation of education funds has no impact on improving the quality of education in Indonesia. Additionally, the budgeted cost for the teacher certification program and school operational assistance absorbs the most the

education funds. A certification that aims to improve the quality of education does not impact teachers' efforts to improve their skills, both in class and on student learning outcomes (Fahmi et al., 2011; Kurniawati et al., 2018; de Ree et al., 2018). The current certification system in Indonesia has no incentive for teachers to improve their performance in the classroom. In fact, the certification allowance provides a financial incentive to earn a bachelor's degree, which is not necessarily proof of being a good teacher (World Bank, 2016).

According to Zulfikar (2010), Indonesian cultural institutions and educational assessment systems play an important role in creating teacher-centered and rote learning in the classroom. Teachers are bound by rules and regulations in a highly centralized top-down instruction system. This makes teachers reluctant to evaluate their instructional pedagogy and tends to teach with a teacher-centered approach. For Indonesian students, teacher support is a strong determinant of their enthusiasm to engage in learning (Maulana et al., 2016). The classroom climate in Indonesia does not show the dialectic characteristic. Classroom climate is only characterized by a teacher-centered approach, where teachers transfer the knowledge to students, and students must memorize and recount during the examinations (Ho et al., 2004). All Initiatives during the learning process in the classroom come from teachers. The ability of students to learn in an autonomous way is not present (Kaluge & Tjahjono, 2004). The contribution of teachers in autonomy support for students was relatively weaker in current Indonesian classroom practice. Therefore, teachers in Indonesia find it difficult to switch to a dialectic approach in the learning climate (Maulana et al., 2016). On the other hand, the relatively low rating of Indonesian teachers on learning climate may also be associated with the still commonly applied student-centered teaching approach (de Ree, 2016; Fasih et al., 2018).

An important aspect is the quality of prospective teachers who will enter and register at public universities to become teachers. In Indonesia, the choice to become a teacher is the second choice and the lowest rated (Suryani et al., 2016). In addition, no special requirements are needed to enroll in a pre-service teacher education program at a public national teacher education institution (Martin, 2019). Perhaps, the reason mentioned above are factors that endorse the low quality of teaching in Indonesia. Teaching is considered a highly skilled career, and with high social status, and is positively correlated with all factors of teacher education (Suryani et al., 2016). Teaching is not just transferring knowledge to students but must have high-level knowledge of skills and have a passion for teaching.

In the Indonesian context, teacher support for student academic engagement is also important. All domains of teaching quality can explain about 45% of the variance in student engagement. Although the level of student engagement was interpreted as moderate, however, it has been proven that student engagement (85%) can be attributed to the class/teacher level (Maulana et al., 2018). It is consistent with past studies originating predominantly western context, in which teacher support for student engagement is important. Teachers in Indonesia have not been fully able to increase student academic engagement. It also contributes to the lower teaching quality in Indonesia.

A safe and stimulating learning climate, classroom management, and clarity of instruction are the basis of quality teaching. Indonesian teachers are severely lacking in these three areas of teaching quality. In fact, the basic skills of teaching quality are skills that must be mastered by novice teachers. Classroom management is important for Indonesian student engagement, its effect seems to be embedded in other domains such as clarity of instruction and teaching-learning strategy (Maulana et al., 2018). We found that actual teaching behavior in terms of classroom management and clarity of instruction is positively correlated with perceived autonomous motivation. Motivational aspects of teaching in the Indonesian education system are not yet explicitly embedded within the curriculum (Iridayanti et al., 2020). Apparently, perceived autonomous motivation is related to the low quality of teaching in Indonesia. In Western countries, such as the Netherlands, classroom management and clarity of teaching are highly emphasized as the first skills that teachers should develop during teacher education. The implementation of realistic teacher education in Netherlands has prioritized classroom management skills to be mastered by novice teachers (van Tartwijk et al., 2011). The lack of basic skills is also one of the causes of the low quality of teaching in Indonesia.

One of the factors measured in this study is teacher motivation. The interaction of teachers and students can determine the success of the learning process in the classroom. Teachers with good teaching behavior will demonstrate effectiveness in teaching, thus leading to good teaching quality as well. The results show that teachers with good teaching effectiveness can increase students' intrinsic motivation in the classroom (Maulana et al., 2016) so that students are motivated to be actively involved in the learning process (Maulana et al., 2015b). This is also supported by research that has been carried out, where the autonomous motivation of teachers in Indonesia can predict the differences in teaching behavior. Evaluation of teaching behavior can be measured by student's engagement in the classroom. The data shows that in general the student's engagement in the classroom is moderate and 85% of student's engagement is determined by the teaching quality of teachers in the classroom.

This finding is related to the Indonesian education system and can be a priority in improving teaching skills which are the responsibility of the Education Personnel Education Institution. We recommend that improvements in teacher motivation, teaching quality profiles and student engagement can contribute to policy recommendations for the Indonesian education system.

6.1 What Needs to Be Improved in the Teaching Quality in Indonesia?

One of the educational problems in Indonesia that must be addressed is the allocation of the education budget. Previously, Indonesia's budget was mostly used for teacher certification programs, and school operational assistance as well as for

teacher incentives. To support the process of improving the quality of teacher education, an effective education budget allocation must be met. Subsequent allocations should be used appropriately to improve the quality of teacher teaching.

Indonesia's main challenge in education is to improve the quality of teacher education. Teacher education institutions must make fundamental changes to improve the teaching quality of the teacher in Indonesia. To achieve that, the requirements for becoming a teacher should be stringent and the standards should be elevated. The teacher professional development must be improved continuously, and it is recommended that periodic evaluations of teacher knowledge and pedagogy should be implemented. The teacher professional development must be designed to address the effective teaching and learning processes in the classroom based on the six domains of teaching and learning.

Furthermore, it is recommended that the workshop and training provided by the government should meet the specific criteria needed by the teachers and give impact on classroom teaching implementation. Training material should be developed to meet the teacher needs based on the classroom observation. The process should be monitored and evaluated periodically to help teachers improve gradually. The certification program should emphasize more on practice and implementation on knowledge and pedagogy and followed by a continuous supervision. Learning from the past failure on certification, teachers are expected to be able to demonstrate their capabilities in the classroom and improve their teaching behavior, not only for one time certification assessment but for continuous progress in the classroom. The most important thing, all of the improvements in teaching quality should give impact to student learning outcomes.

Based on the factors that contribute to Indonesia low teaching quality, the teacher's lack of content knowledge, we suggest the result of our study give insight on what to do to improve Indonesia teaching quality. Our study focuses on the process in the classroom and the interaction between teacher and students. The six domains of teaching learning behavior can be used as a benchmark for teacher quality improvement in the classroom. By improving the teacher competencies in the six domains of teaching behavior, also give chances to increase student's engagement.

It can be concluded that in general, the profile of teaching quality in Indonesia is still relatively low based on both observer perception and student perception. In all domain's effective teaching behavior is moderate/sufficient, except for the differentiation instruction domain is low/insufficient. Meanwhile, all domains of teaching behavior seen by student perception in Indonesia was categorized as moderate/sufficient. These findings a strong basis for Indonesian teachers to improve their teaching behavior, especially in domain adaptation of teaching/differentiation and maybe also for the other domains.

Acknowledgements We would like to express our gratitude to all Indonesian observers and teachers who participated in this study. This work was cooperation between Indonesia and the Netherlands and partially supported by the Dutch scientific funding agency (NRO, project number: 405-15-732) and the Directorate General of Higher Education fund of Indonesia (project number: SK No.12/SP2H/DRPM/LPPM-UNJ/III/2019). Parts of the present study were presented during the ISATT 2017 conference in Salamanca, Spain.

References

- André, S., Maulana, R., Helms-Lorenz, M., Telli, S., Chun, S., Fernández-García, C. M., et al. (2020). Student perceptions in measuring teaching behavior across six countries: A multi-group confirmatory factor analysis approach to measurement invariance. *Frontiers in Psychology, 11*, 273. <https://doi.org/10.3389/fpsyg.2020.00273>
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133–180. <https://doi.org/10.3102/0002831209345157>
- Beijaard, D., Korthagen, F., & Verloop, N. (2007). Understanding how teachers learn as a prerequisite for promoting teacher learning. *Teachers and Teaching, Theory and Practice, 13*(2), 105–108. <https://doi.org/10.1080/13540600601152298>
- Ben-Peretz, M. (1980). Teachers' role in curriculum development: An alternative approach. *Canadian Journal of Education/Revue canadienne de l'éducation, 52*–62. <https://doi.org/10.2307/1494313>
- Benton, S. L., & Cashin, W. E. (2012). Idea paper# 50 student ratings of teaching: A summary of research and literature.
- Blömeke, S., & Delaney, S. (2014). Assessment of teacher knowledge across countries: A review of the state of research. *International Perspectives on Teacher Knowledge, Beliefs, and Opportunities to Learn, 541*–585. https://doi.org/10.1007/978-94-007-6437-8_25
- Central Intelligence Agency. (2017). *The World Factbook*. Retrieved February 4, 2017, from <https://www.cia.gov/library/publications/the-world-factbook/geos/id.html>
- De Jong, R., & Westerhof, K. J. (2001). The quality of student ratings of teacher behaviour. *Learning Environments Research, 4*(1), 51–85. <https://doi.org/10.1023/A:1011402608575>
- De Ree, J. J. (2016). *Indonesia-teacher certification and beyond: An empirical evaluation of the teacher certification program and education quality improvements in Indonesia* (No. 104599, pp. 1–76). The World Bank.
- De Ree, J., Muralidharan, K., Pradhan, M., & Rogers, H. (2018). Double for nothing? Experimental evidence on an unconditional teacher salary increase in Indonesia. *The Quarterly Journal of Economics, 133*(2), 993–1039. <https://doi.org/10.1093/qje/qjx040>
- Den Brok, P., Brekelmans, M., & Wubbels, T. (2004). Interpersonal teacher behaviour and student outcomes. *School Effectiveness and School Improvement, 15*(3–4), 407–442. <https://doi.org/10.1080/09243450512331383262>
- Den Brok, P., Brekelmans, M., & Wubbels, T. (2006). Multilevel issues in research using students' perceptions of learning environments: The case of the questionnaire on teacher interaction. *Learning Environments Research, 9*(3), 199–213. <https://doi.org/10.1007/s10984-006-9013-9>
- Downer, J. T., Rimm-Kaufman, S. E., & Pianta, R. C. (2007). How do classroom conditions and children's risk for school problems contribute to children's behavioral engagement in learning? *School Psychology Review, 36*(3), 413–432. <https://doi.org/10.1080/02796015.2007.12087938>
- Fahmi, M., Maulana, A., & Yusuf, A. A. (2011). Teacher certification in Indonesia: A confusion of means and ends. *Center for Economics and Development Studies (CEDS) Padjadjaran University, 3*(1), 1–18. <http://www.ceds.fe.unpad.ac.id/> and <http://econpapers.repec.org/paper/unpwpaper/>
- Fasih, T., Afkar, R., & Tomlinson, H. B. (2018). *Learning for all: Towards quality education for enhanced productivity and economic growth in Indonesia* (No. 123652, pp. 1–28). The World Bank.
- Fend, H. (2002). Mikro-und Makrofaktoren eines Angebot-Nutzungsmodells von Schulleistungen (micro-and macrofactors of an offer-usage model of school achievement). *Zeitschrift für Pädagogische Psychologie, 15*, 141–149.
- Gagne, R. M., & Briggs, L. J. (1974). *Principles of instructional design*. Holt, Rinehart & Winston.
- Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. National comprehensive centre for teacher quality.
- Goe, L., Bell, C., & Little, O. (2008). Approaches to evaluating teacher effectiveness: A research synthesis. *National Comprehensive Center for Teacher Quality*.
- Goldhaber, D. (2015). Teacher effectiveness research and the evolution of US teacher policy. The productivity for results series no. 5. George W. Bush Institute, Education Reform Initiative.

- Hambleton, R. K., Merenda, P., & Spielberger, C. (Eds.). (2004). *Adapting educational and psychological tests for cross-cultural assessment*. Lawrence Erlbaum Publishers.
- Hamid, S. R., Syed Hassan, S. S., & Ismail, N. A. H. (2012). Teaching quality and performance among experienced teachers in Malaysia. *Australian Journal of Teacher Education*, 37(11), 5.
- Harrell, P., Leavell, A., van Tassel, F., & McKee, K. (2004). No teacher left behind: Results of a five-year study of teacher attrition. *Action in Teacher Education*, 26(2), 47–59. <https://doi.org/10.1080/01626620.2004.10463323>
- Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher*, 41(2), 56–64. <https://doi.org/10.3102/0013189X12437203>
- Ho, E., Holmes, P., & Cooper, J. (2004). *Review and evaluation of the literature on managing cultural diversity in the classroom. Report for the Ministry of Education and Education New Zealand*. The University of Waikato.
- Houtveen, A. A. M., & Van de Grift, W. J. C. M. (2007). Effects of metacognitive strategy instruction and instruction time on reading comprehension. *School Effectiveness and School Improvement*, 18(2), 173–190. <https://doi.org/10.1080/09243450601058717>
- Irmidayanti, Y., & Fadhilah, N. (2018, January). Indonesian secondary education teachers' teaching behavior: General profiles and relation to student engagement. In R. Maulana (Chair). *Effective teaching behavior and student engagement in secondary education: A multi-national symposium*. Symposium conducted at the International Congress for School Effectiveness and Improvement, Singapore.
- Irmidayanti, Y., Maulana, R., Helms-Lorenz, M., & Fadhilah, N. (2020). Relationship between teaching motivation and teaching behaviour of secondary education teachers in Indonesia (Relación entre la motivación docente y el comportamiento docente en profesores de educación secundaria en Indonesia). *Journal for the Study of Education and Development*, 43(2), 271–308. <https://doi.org/10.1080/02103702.2020.1722413>
- Jalal, F., Samani, M., Chang, M. C., Stevenson, R., Ragatz, A. B., & Negara, S. D. (2009). *Teacher certification: A strategy for teacher quality improvement*. The Ministry of National Education and The World Bank.
- Kaluge, L., & Tjahjono, H. (2004). *The quality improvement of primary children learning through a school-based programme in Indonesia*. University of Surabaya.
- Kemendikbud. (2016). *Program sertifikasi*. <https://www.kemendikbud.go.id/main/blog/2016/04/kemendikbud-aniesbaswedan-program-sertifikasi>
- Kunter, M., Baumert, J., & Köller, O. (2007). Effective classroom management and the development of subject-related interest. *Learning and Instruction*, 17(5), 494–509. <https://doi.org/10.1016/j.learninstruc.2007.09.002>
- Kunter, M., Tsai, Y. M., Klusmann, U., Brunner, M., Krauss, S., & Baumert, J. (2008). Students' and mathematics teachers' perceptions of teacher enthusiasm and instruction. *Learning and Instruction*, 18(5), 468–482. <https://doi.org/10.1016/j.learninstruc.2008.06.008>
- Kurniawati, S., Suryadarma, D., Bima, L., & Yusrina, A. (2018). Education in Indonesia: A white elephant? *Journal of Southeast Asian Economies*, 35(2), 185–199. <https://www.jstor.org/stable/26539213>
- Kyriakides, L. (2005). Drawing from teacher effectiveness research and research into teacher interpersonal behaviour to establish a teacher evaluation system: A study on the use of student ratings to evaluate teacher behaviour. *The Journal of Classroom Interaction*, 44–66. <https://www.jstor.org/stable/23870663>
- Kyriakides, L., Creemers, B. P., & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education*, 25(1), 12–23. <https://doi.org/10.1016/j.tate.2008.06.001>
- Lawrence-Brown, D. (2004). Differentiated instruction: Inclusive strategies for standards-based learning that benefits the whole class. *American Secondary Education*, 32(3), 34–62.
- Lawrenz, F., Huffman, D., & Robey, J. (2003). Relationships among student, teacher and observer perceptions of science classrooms and student achievement. *International Journal of Science Education*, 25(3), 409–420. <https://doi.org/10.1080/09500690210145800>
- Martin, S. N. (2019). Science education in Indonesia: Past, present, and future. *Asia-Pacific Science Education*, 5(1), 1–29. <https://doi.org/10.1186/s41029-019-0032-0>

- Martin, M. O., Mullis, I. V. S., & Hooper, M. (Eds.). (2016). Methods and procedures in TIMSS 2015. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timssandpirls.bc.edu/publications/timss/2015-methods.html>
- Martínez-Rizo, F. (2012). Procedimientos para el estudio de las prácticas docentes. Revisión de la literatura. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, 18(1), 1–22.
- Maulana, R., & Helms-Lorenz, M. (2016). Observations and student perceptions of the quality of preservice teachers' teaching behaviour: Construct representation and predictive quality. *Learning Environments Research*, 19(3), 335–357. <https://doi.org/10.1007/s10984-016-9215-8>
- Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2015a). Development and evaluation of a questionnaire measuring pre-service teachers' teaching behaviour: A Rasch modelling approach. *School Effectiveness and School Improvement*, 26(2), 169–194. <https://doi.org/10.1080/09243453.2014.939198>
- Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2015b). Pupils' perceptions of teaching behaviour: Evaluation of an instrument and importance for academic motivation in Indonesian secondary education. *International Journal of Educational Research*, 69, 98–112. <https://doi.org/10.1016/j.ijer.2014.11.002>
- Maulana, R., Helms-Lorenz, M., Irnidayanti, Y., & van de Grift, W. (2016). Autonomous motivation in the Indonesian classroom: Relationship with teacher support through the lens of self-determination theory. *The Asia-Pacific Education Researcher*, 25(3), 441–451. <https://doi.org/10.1007/s40299-016-0282-5>
- Maulana, R., Helms-Lorenz, M., & Irnidayanti, Y. (2017a). Teaching quality of Indonesian secondary school teachers: Profile, explanatory variables, and importance for student engagement. In *International study association on teachers and teaching (ISATT)*. Salamanca, .
- Maulana, R., Helms-Lorenz, M., & Van de Grift, W. (2017b). Validating a model of effective teaching behaviour of pre-service teachers. *Teachers and Teaching*, 23(4), 471–493. <https://doi.org/10.1080/13540602.2016.1211102>
- Maulana, R., Irnidayanti, Y., Fokkens-Bruinsma, M., & Helms-Lorenz, M. (2018). Indonesian students' academic engagement and the role of teachers' teaching behavior in econdary education. In *Asian education miracles* (pp. 63–83). Routledge.
- Maulana, R., André, S., Helms-Lorenz, M., Ko, J., Chun, S., Shahzad, A., et al. (2020). Observed teaching behaviour in secondary education across six countries: Measurement invariance and indication of cross-national variations. *School Effectiveness and School Improvement*, 32(1), 64–95.
- Montague, M., & Dietz, S. (2009). Evaluating the evidence base for cognitive strategy instruction and mathematical problem solving. *Exceptional Children*, 75(3), 285–302. <https://doi.org/10.1177/001440290907500302>
- Montuoro, P., & Lewis, R. (2014). Student perceptions of misbehavior and classroom management. In *Handbook of classroom management* (pp. 354–372). Routledge.
- Muijs, D., Kyriakides, L., Van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art—teacher effectiveness and professional learning. *School Effectiveness and School Improvement*, 25(2), 231–256. <https://doi.org/10.1080/09243453.2014.885451>
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). TIMSS 2015 international results in mathematics. Retrieved from Boston College, TIMSS & PIRLS International Study Center. website: <http://timssandpirls.bc.edu/timss2015/international-results/>
- OECD/Asian Development Bank. (2015). *Education in Indonesia: Rising to the challenge*. OECD Publishing. <https://doi.org/10.1787/9789264230750-en>
- Pearson, P. D., & Fielding, L. (1991). Comprehension instruction. In R. Barr, M. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research*, 2 (pp. 815–860). Longman.
- Peterson, K. D., Wahlquist, C., & Bone, K. (2000). Student surveys for school teacher evaluation. *Journal of Personnel Evaluation in Education*, 14(2), 135–153. <https://doi.org/10.1023/A:1008102519702>
- PISA. (2015). *Assessment and analytical framework science, reading, mathematics, financial literacy and collaborative problem solving*. OECD Publications Service.
- Praetorius, A. K., Lenske, G., & Helmke, A. (2012). Observer ratings of instructional quality: Do they fulfill what they promise? *Learning and Instruction*, 22(6), 387–400. <https://doi.org/10.1016/j.learninstruc.2012.03.002>

- Pressley, M., Woloshyn, V., Lysynchuk, L. M., Martin, V., Wood, E., & Willoughby, T. (1990). A primer of research on cognitive strategy instruction: The important issues and how to address them. *Educational Psychology Review*, 2(1), 1–58. <https://doi.org/10.1007/BF01323528>
- Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Economic Policy Institute, 1660 L Street, NW, Suite 1200, Washington, DC 20035.
- Richardson, J. T. (2005). Instruments for obtaining student feedback: A review of the literature. *Assessment & Evaluation in Higher Education*, 30(4), 387–415. <https://doi.org/10.1080/02602930500099193>
- Roberts, M. J., & Erdos, G. (1993). Strategy selection and metacognition. *Educational Psychology*, 13(3–4), 259–266. <https://doi.org/10.1080/0144341930130304>
- Rotgans, J. I., & Schmidt, H. G. (2011). The role of teachers in facilitating situational interest in an active-learning classroom. *Teaching and Teacher Education*, 27(1), 37–42. <https://doi.org/10.1016/j.tate.2010.06.025>
- Scherer, R., Nilsen, T., & Jansen, M. (2016). Evaluating individual students' perceptions of instructional quality: An investigation of their factor structure, measurement invariance, and relations to educational outcomes. *Frontiers in Psychology*, 7, 110. <https://doi.org/10.3389/fpsyg.2016.00110>
- Scherzinger, M., & Wettstein, A. (2019). Classroom disruptions, the teacher–student relationship and classroom management from the perspective of teachers, students and external observers: A multimethod approach. *Learning Environments Research*, 22(1), 101–116. <https://doi.org/10.1007/s10984-018-9269-x>
- Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77(4), 454–499. <https://doi.org/10.3102/0034654307310317>
- Subban, P. (2006). Differentiated instruction: A research basis. *International Education Journal*, 7(7), 935–947.
- Suryani, A., Watt, H. M., & Richardson, P. W. (2016). Students' motivations to become teachers: FIT-choice findings from Indonesia. *International Journal of Quantitative Research in Education*, 3(3), 179–203.
- TIMSS. (2015). *International results in mathematics*. Retrieved from Boston College, TIMSS & PIRLS International Study Center. <http://timssandpirls.bc.edu/timss2015/international-results/>
- Todd, A., & Mason, M. (2005). Enhancing learning in south African schools: Strategies beyond outcomes-based education. *International Journal of Educational Development*, 25(3), 221–235. <https://doi.org/10.1016/j.ijedudev.2004.08.003>
- Tomlinson, C. A. (1999). *The differentiated classroom, responding to the needs of all learners, association for supervision and curriculum development*. Alexandria.
- Tomlinson, C. A. (2005). Traveling the road to differentiation in staff development. *Journal of Staff Development*, 26(4), 8–12.
- van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research*, 49(2), 127–152. <https://doi.org/10.1080/00131880701369651>
- van de Grift, W., Helms-Lorenz, M., & Maulana, R. (2014). Teaching skills of student teachers: Calibration of an evaluation instrument and its value in predicting student academic engagement. *Studies in Educational Evaluation*, 43, 150–159. <https://doi.org/10.1016/j.stueduc.2014.09.003>
- van der Lans, R. M., van de Grift, W. J., & van Veen, K. (2015). Developing a teacher evaluation instrument to provide formative feedback using student ratings of teaching acts. *Educational Measurement: Issues and Practice*, 34(3), 18–27. <https://doi.org/10.1111/emip.12078>
- van Tartwijk, J., Veldman, I., & Verloop, N. (2011). Classroom management in a Dutch teacher education program: A realistic approach. *Teaching Education*, 22(2), 169–184. <https://doi.org/10.1080/10476210.2011.567847>
- Vandeyar, S., & Killen, R. (2007). Educators' conceptions and practice of classroom assessments in post-apartheid South Africa. *South African Journal of Education*, 27(1), 101–115.
- Vermunt, J. D., & Verloop, N. (1999). Congruence and friction between learning and teaching. *Learning and Instruction*, 9(3), 257–280. [https://doi.org/10.1016/S0959-4752\(98\)00028-0](https://doi.org/10.1016/S0959-4752(98)00028-0)

- Wagner, W., Göllner, R., Werth, S., Voss, T., Schmitz, B., & Trautwein, U. (2016). Student and teacher ratings of instructional quality: Consistency of ratings over time, agreement, and predictive power. *Journal of Educational Psychology*, 108(5), 705–721.
- Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63(3), 249–294. <https://doi.org/10.3102/0346543063003249>
- Wei, M., Den Brok, P., & Zhou, Y. (2009). Teacher interpersonal behaviour and student achievement in English as a foreign language classrooms in China. *Learning Environments Research*, 12(3), 157–174. <https://doi.org/10.1007/s10984-009-9059-6>
- World Bank (IBRD). (2016). *Teacher certification and beyond: An empirical evaluation of the teacher certification program and education quality improvements in Indonesia*. <http://hdl.voced.edu.au/10707/429757>
- Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (1997). *Program evaluation: Alternative approaches and practical guidelines* (2nd ed.). Longman.
- Zulfikar, T. (2010). The making of Indonesian education: An overview on empowering Indonesian teachers. *Journal of Indonesian Social Sciences and Humanities*, 2, 13–39. <http://www.kitlv-journals.nl/index.php/jissh/index>

Yulia Irnidayanti obtained her first degree in Biology Education and PhD in Biology. She is currently a Senior Lecturer and researcher at the Biology and Biology Education Department, Universitas Negeri Jakarta [State University of Jakarta], Indonesia. Since 2001, she has been working together with the Teacher Education Department of University of Groningen, the Netherlands, on the project about teaching quality and student academic motivation from the international perspective (ICALT3/Differentiation project, Principal investigator Indonesia). She is interested in helping teachers to improve their teaching quality and student differences in their learning needs, motivation, and learning style.

Nurul Fadhilah is a university lecturer at the Department of Biostatistic and Population, University of Indonesia. She has been actively involved in the international project called ICALT3/Differentiation as an expert observer and as co-investigator for Indonesia. She is currently involved in a research project involving public health big data analysis. She has been involved in professional teacher development for high school teachers in DKI Jakarta. She is experienced in designing and facilitating teacher professional development training, developing syllabus, task designing, developing differentiated instructions, especially in Cambridge IGCSE and A level Biology subject.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 11

Effective Teaching in Mongolia: Policies, Practices and Challenges



Amarjargal Adiyasuren and Ulziisaikhan Galindev

Abstract This chapter describes the contextual background of teacher and teaching quality in Mongolia through exploring teacher policies, and practices and challenges surrounding the teacher, followed by how curriculum sets the parameters for teaching behaviour. Students must finish a four-year teacher education program in Mongolia to become teachers. The government policy aims to increase the percentage of teachers who hold master's degrees up to 70% by 2024; 15.8% of primary and secondary education teachers held a master's degree as of 2020. The government requires teachers to attend mandatory training in their first, fifth and tenth teaching year. Besides these centralized trainings, the government is also reinforcing teachers' professional development policies in the direction that supports and encourages local and school-based professional development based on teachers' learning needs. Recently there has been a regulation of school self-monitoring and evaluation, including setting criteria on lesson management and quality to use for evaluation of teachers' teaching skills and behaviour, via lesson observations. Teacher behaviour and pedagogical methods are articulated in the curriculum documents as well. The most recent education reform was aimed at a principle that is called the change of 'Each and every child'. This was followed by curriculum revision with key concepts of inquiry-based learning, differentiating teaching (based on students' developmental differences) and assessment of progress and learning skills. These changes, needless to say, require teachers to improve their pedagogical skills. Research shows that Mongolian teachers still have difficulty with devising differentiated activities for students at different levels of learning. In terms of context, it should be understood that teaching is regarded as a low paid profession in Mongolia. The government takes measures such as: offering scholarships to attract good students into teaching profession; and providing salary supplements and local subsidies.

Keywords Mongolia · Teacher policies · Teacher quality · Differentiating instruction · Teaching learning strategies

A. Adiyasuren (✉) · U. Galindev

The Department of Educational Administration, Mongolian National University of Education, Ulaanbaatar, Mongolia

e-mail: a.amarjargal@gmail.com; ulziisaikhan@msue.edu.mn

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_11

245

1 Introduction

There have been major reforms in all levels of education in Mongolia since the collapse of the socialist system in 1990. As the fundamental social value shifted to democratic and humanistic philosophy, education systems including curriculum, content, pedagogy and governance, needed to shift. UNESCO (2019) remarked on the great effort of teachers who had overcome the challenges of past decades and brought the education system up to date.

Education systems compare their quality of education and student achievement through international benchmarking studies such as the Programme for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS). Mongolia is planning to participate in PISA for the first time in 2022. Although Mongolia attended TIMSS in 2007, the achievement result was excluded from the comparison because of poor documentation of the samples and data. Accordingly, there is no data available about the education system in terms of student performance/achievement through international comparative studies. A recent study says teacher quality in Mongolia, in terms of policies and mechanisms, is higher than average than some other Asian countries (Chun & Gentile, 2020). Within the country, research on teaching quality and behaviour are scarce.

It is important to understand the contextual background of the quality of teachers and teaching in Mongolia. We pose the following research question and sub questions.

Research question: What is the contextual background for quality of teachers and teaching in Mongolia?

Sub research question 1: What are the teacher policies and challenges around them?

Sub research question 2: What are some curriculum related factors that guide teachers' teaching skills and behaviour?

2 Policies and Challenges

In this section, the current system, including teacher related policies and mechanisms, from initial teacher preparation to entry to teaching profession, including professional development and related factors will be described.

2.1 *Teacher Preparation*

Primary and secondary education teacher training is offered as a four-year bachelor of education course of study. Graduates of secondary education teacher programs are qualified to teach both at lower (grade 6–9) and upper secondary (grade 10–12) level. Secondary education teachers of all subjects teach grades 6–12.

More than half of school and kindergarten teachers study for their qualifications at the Mongolian National University of Education. In 2020 it was recorded that 7.8% of the teachers hold a diploma education, 76% of the teachers hold a bachelor's degree and 15.8% hold a master's degree or above. The government policy aims to increase the percentage of teachers who hold master's degrees to 70% by 2024, similarly to some developed countries such as Finland.

Criticism of teaching quality tends to focus on teacher education (Gore et al., 2001). Mongolia is not an exception. As this profession is regarded as a low paid profession in Mongolia, pre-service teacher candidates tend to have lower university entrance examination scores than the other specialisations. In 2013, the government started offering a scholarship to those with high university entrance examination scores who wished to pursue the teaching profession in order to encourage high calibre candidates. This program became a crucial measure for increasing the quality of candidates enrolling for pre-service courses (UNESCO, 2019). Although the idea was good and attracted many students, neither the ministry nor the university ensured that graduates would choose to enter to the teaching positions after graduation.

The current teacher standards were approved in 2010. These standards covered teacher training program curriculums, evaluations, duration, and requirements for learning environment. The teacher standards clearly stated the necessary competences and behaviors of the graduates of the teacher major. The standard also includes mandatory and elective courses and a minimum number of credits. However, the standard is not consistently implemented, and teacher training universities and programs lack a comprehensive policy and a consolidated curriculum. This leads to a system which produces a variety of teachers, including some who are poorly prepared and un-qualified.

It is important to consider that there is no official support system or induction program for novice teachers at school. This fact makes the quality of teacher training even more important. In a small survey, teachers answered that 88% of their teaching knowledge was learnt in teacher training and 83% answered that the program was good or very good (Enkhtuvshin, 2020). For teacher training subject content, pedagogy and teaching practice are the most important elements for quality teacher preparation. 77% of the teachers answered that they were prepared well or very well.

2.2 Entry to the Profession

School principals hold the full authority to hire and allocate teachers in Mongolia. For new graduates, a Bachelor degree from a teacher education program is considered a teaching license.

Urban schools and local schools face different problems regarding hiring teachers. Once pre-service candidates attend a capital city for university, they like to stay in the city as teachers. Overall, well-educated, skilled and experienced primary and

secondary education teachers, specifically Mathematics, English, Physics and ICT teachers are unwilling to work in rural areas where laboratories, teaching aids and other resources are in scarce supply (UNESCO, 2019). Rural schools are always in need of teachers. Some rural schools offer accommodation to attract new graduate or young teachers who are in need of financial support. The government also provides additional ‘local subsidy’ every five consecutive years to keep teachers at rural schools.

In 2014, a regulation was introduced that required new graduates to qualify to become novice teachers. However, the teacher qualification examination was withdrawn in 2018, because there was very low interest from candidates to enter the teaching profession in rural areas. Only around forty percent of the exam takers passed (UNESCO, 2019). The system was not equipped to verify the teachers’ ability to practice in the field (Kim et al., 2017). The researchers also found that the system was criticized for failing to guarantee conformity, fairness and transparency; however a policy review suggests that the teacher qualification examination should return from a legal standpoint (UNESCO, 2019).

Although teachers are guaranteed to life-long job and stable economic rewards as government officials, teachers’ social recognition is low in Mongolia.

2.3 Teacher Professional Development

2.3.1 National Level Professional Development

The Mongolian system of in-service teacher education was similar to that of the Soviet Union and other socialist countries, when established in 1969 (Steiner-Khamsi, 2005). A prominent feature of the socialist system was “life-long learning” which included the right of each teacher and administrator to attend centrally organized teacher education sessions every five years.

Around 2000, the focus of national and donor-driven teacher quality reform activities were directed to improvements to the in-service teacher training system, leaving the pre-service system neglected and under-funded (ADB, 2008). In-service development programs needed to fill the gaps in knowledge and skills of teachers which should have been inculcated during pre-service teacher training. Needs-based, decentralized in-service teacher training was implemented through the ‘Voucher system’, adopted in 1998. It was intended to allow schools to choose the type of teacher training based on school and teacher needs (Pagma et al., 2002), but the practice was not effective. Teachers, school principals and provincial education authorities abused the vouchers for visiting the capital city (Steiner-Khamsi & Stolpe, 2006).

Shifted back to the centralized professional development system, a ministry-affiliated Institute for Teachers’ Professional Development (ITPD) was re-established in 2012 and offered centralized mandatory training to teachers in their

first, fifth and tenth year of teaching. For some years the training functioned as an extension of their teaching license, but then was withdrawn. The focus of the training was ‘learning’, ‘collaborating’, and ‘sharing knowledge and experience’. All the expenses related to mandatory training are paid for by the government. Centralized training for 40 h consisted of 4 h of policy and legal training, 4 h personal development, 8 h of ICT skills, and 22 h of professional knowledge and methodology. The system provides teachers an equal opportunity to improve teachers’ knowledge, methodology and skills which is important in terms of equality.

2.3.2 Local and School Level Professional Development

The most recent regulation ‘Promoting teacher development law’ of 2018 encouraged the decentralization of teacher professional development. Even though the centralized training remained the same, to ensure equal opportunity for teachers and local units (provinces and districts in the capital level) schools were required to establish ‘Teacher development centers’ for teachers to develop their knowledge and skills sustainably on the job.

Even though local level education departments provide the teachers with opportunities to share their knowledge and experience, the practice varies depending on the initiatives of the officials in local education department. School level supervision is organized by subject-based teacher groups in secondary level and grade-based teacher groups in primary level. Teacher induction programs for novice teachers are very poor at schools.

The teacher promotion system is based on professional degrees: regular teacher, methodologist teacher, leading teacher, advisor teacher. Teachers are expected to aim to get the degree when the working year requirement was fulfilled because the promotion criteria was mainly based on working years up until 2018. With the new law of 2018, general requirements of student learning achievement, teacher professional and methodological skill, satisfaction of learners, teachers and peers, parents and caretakers, self-development have to be fulfilled in order to promote to a next degree.

2.3.3 Teacher Evaluation, Appraisal and Salary

Teachers who themselves, or whose students, successfully attend academic competitions were considered “good teachers” in the past. School evaluation and teacher evaluation both included criteria such as the preparation of students for national or international academic competitions such as International Mathematical Olympiad, their participation, and their performance. Competition achievement was tied with teacher performance and salary system. Teacher’s salary consisted of a base salary, supplement salary and bonuses. The base salary of teachers was solely based on a teacher’s experience. Supplement salary for teachers was introduced in Mongolia in 1995 (World Bank, 2006). Supplement salary was provided based on being a

homeroom teacher, incentives for overtime, remuneration for teacher's professional degree, taking charge of cabinet or laboratory, leading the subject teaching sector, remuneration for the skills, or remuneration for residing in rural areas.

When outcomes-based education was introduced to Mongolia in 2003, teacher salaries were tied to performance and teacher 'outcome contract' or scorecard as it was called. The teacher performance requirement included 10 criteria/indicators, only two of which were directly linked to students, class management and student development. Bonuses were four-time awards given once a year, based on evaluation by the school administration of the teacher performance.

Continuing to the current system, the education reform of 2012 emphasized "developing each and every student" and this changed the concept of a 'skilled' or 'good' teacher. Criticism that 'teachers only focus on national competition promising students and ignore the rest' changed the requirements. The current teacher evaluation system assesses teachers' performance by five criteria which include: students' academic achievement, character development, talent, health, and parents' satisfaction. Quarterly incentive supplement bonuses are based on the result of both a teacher self-evaluation and an evaluation of the school principal or instructional manager based on the five criteria. In a recent study about school management, more than 70% of teachers answered teacher evaluation conducted by school management help teachers to improve their lesson (ADB, 2017). However, the common practice is that a school's total amount for incentive supplement is divided equally across all teachers regardless of individual teachers' performance. The salary supplements account for around 41% of a teacher's income (UNESCO, 2019) so this is a critical issue. It should also be noted that the salary supplement received for teaching additional hours makes up the largest percentage of a teacher's monthly income excluding the base salary.

As a mechanism for teachers to be recognized and rewarded for their teaching, the government is planning to introduce a performance-based salary system in the near future. Prior to this reform, the government has approved a new school self-monitoring and evaluation regulation in 2019 that includes evaluation rubrics with five domains to evaluate the school; one of these domains pertains to lesson management and quality, which is directly linked to teachers' teaching behavior. The domain consists of 17 items which the school principal or education manager must monitor in order to evaluate teachers' teaching through observations. It can be expected that the observations would be used as useful data for teacher professional development and improving their teaching in the classroom. Interestingly, some criteria of teachers' teaching behavior included in the regulation look very similar to some "International Comparative Analysis of Teaching and Learning" (ICALT, will be explained in next section) items: 5 items to safe and stimulating educational climate, 3 items to clear and structured instruction, 1 items to teaching learning strategies, and 2 items to differentiating instruction.

The TALIS 2013 questionnaire reveals the job satisfaction of Mongolian teachers. With a 4-point rating scale, the average job satisfaction of teachers was 3.42 (Ulziisaikhan, 2017). Job satisfaction about work environment was 3.33. By working experience, teachers up to 5 years and over 21 years have the highest job

satisfaction, which is the same as the TALIS-2013 result. The study showed that it was not a school related factor but the teacher-student relationship and teacher collaboration which was a positive factor in their job satisfaction.

3 Teachers' Teaching Skill and Behavior

Questions arise about what the level of the actual teaching skill and behaviour is in the classroom. This section answers sub research question 2: What are some curriculum related factors that guide teachers' teaching skill and behavior?

Although there are studies about quality of teaching of Mongolian teachers (Jadamba et al., 2014; Enkhtuvshin, 2014; Luvsandorj & Oyun-Erdene, 2015), none of this research includes actual teaching in classroom. In a study that measured and compared the effective teaching behavior of teachers of Mongolia and Korea (Chun et al., 2020), Korean teachers performed higher than Mongolian teachers, although with a small difference. The theoretical framework and tool called "International Comparative Analysis of Learning and Teaching (ICALT)" can be used to study effective teaching behaviour, which includes six domains. Safe and stimulating educational climate, Efficient classroom management, Clear and structured instruction, Intensive and activating teaching, Teaching learning strategies, and Differentiating instruction (van de Grift et al., 2014). ICALT tool measures teaching behaviour by these six domains through 35 items, using four ordinal response categories (1 = 'mostly weak' to 4 = 'mostly strong'). Of these six domains, the first three domains refer to basic teaching skills, and the second three refer to advanced teaching skills.

The comparative analysis of the teaching quality of Mongolian and Korean secondary teachers using ICALT verified the tools and the feasibility of comparing teaching quality (Chun et al., 2020). Mongolian teachers were rated 3.20 in average in safe and stimulating educational climate domain, 3.03 in efficient classroom management domain, 2.95 in clear and structured instruction, and lower than 2.7 in all advanced teaching skill domains including the lowest, 2.41 in differentiating instruction domain (Chun et al., 2020).

We assumed that the teaching behaviors articulated in the ICALT observation tool align with the direction of educational reforms in Mongolia. And we investigated the following questions:

- What are research findings and practices regarding teaching skill and behaviour?
- How does the national curriculum guide teaching behaviour?

Mongolia has been implementing education reforms based on learning from international systems and experiences for the last three decades. As mentioned in the previous section, teachers who prepared and successfully sent their students to subject competitions were considered 'good teachers' in the socialist period and after that. A new government established in 2012 initiated a program "Upright Mongolian

child” that brought primary and secondary education reform. A criticism at the time was that teachers had focused on strong students with potential and left behind the mass. The concept of “developing each and every child” led teachers to work in different ways. Subject competitions for primary education were prohibited and many schools stopped providing subject intensive programs that were targeted for competitions. Instead, more inclusive principles such as providing equal opportunities for every student, referring to students’ developmental differences, developing each student’s talent, interest and characteristics and lastly, equipping students with learning strategies were strongly required from schools and teachers. Educational goals and objectives integrated more twenty-first century skills and give more emphasis on learning skills in primary and secondary education.

In particular, the reform is aligned with teachers’ skill and behaviour as articulated in the advanced skills of ICALT tool, Differentiating instruction and Teaching learning strategies domains, and defines the teachers’ skill and behaviour in some extent.

3.1 Differentiating Instruction

Differentiating instruction in the classroom has been encouraged strongly for the last 10 years. Integrating differentiated instruction principles and practices and providing differentiated learning tasks according to students’ ability or learning levels in daily classroom practice was introduced through Mongolia-Cambridge Education Initiative, a curriculum reform prior to 2012-year reform. Formative assessment was also another new strategy systematically introduced to Mongolian teachers with the Mongolia-Cambridge Education Initiative.

An increasing humanistic view in society is also affecting education systems in Mongolia in terms of differentiation. The inclusive education agenda has regulated that up to one or two students with special needs can learn in each class. Teachers are expected to gain wider and deeper knowledge and methods of inclusive education including differentiated instruction. Inclusive education has become one of the mandatory programs in centralized in-service teacher training.

Differentiating instruction is a complex thing. Most teachers admit that they need professional development to devise differentiated activities for different level of learners and new strategies on classroom management (ADB, 2017). In research on the implementation of curriculum, 40% or more of teachers want more professional development training in the areas of how to teach the new curriculum, update their knowledge and understanding of their specialist field, improve their pedagogical skills, formative and summative assessment, classroom management and individualizing learning as well as catering to learners with special needs (ADB, 2017).

3.2 *Teaching Learning Strategies*

A major objective of the introduction of the new curriculum was to increase student learning outcomes through better learning strategies. The National core curriculum document not only shows the content area, but also provides the pedagogies for per subject through the learning objectives. For Mathematics and Social Science, the learning objectives are defined with in a problem-solving learning paradigm, for Science in inquiry-based learning, Mongolian language in Information processing and Design and technology in Project-based or product-based learning. Teachers need to acquire new skills and teaching expertise accordingly.

Moreover, along with the curriculum reform, Mongolia has adopted and adapted the student learning evaluation system from Japan (*'kantenbetsu'* evaluation system). Student learning evaluation system consists of three aspects that are knowledge and understanding, learning skill, and attitude.

However, a remaining problem is that teachers do not understand what the learning skills or learning strategies look like in the classroom. Not having themselves learnt in this way, they do not know how to teach in this way (ADB, 2017).

4 Conclusion and Discussion

In this chapter we explained the contextual background of teacher and teaching quality in Mongolia by reviewing the policies, some practices and challenges.

A recent study says teacher quality of Mongolia in terms of policies and mechanisms is above the average in Asian countries (Chun & Gentile, 2020) and our analysis does reveal some good policies. However, the policy coherence linking teacher preparation, teacher professional development, and teachers' evaluation appears weak, and some policies are not being implemented sufficiently in all settings.

Teacher professional standards and government scholarship attract the best students into the teaching profession and there are good policies and practices in initial teacher preparation. However, policy implementation is ignored or not monitored by those who should be responsible.

Professional development systems have been changed several times in the last 10 years. The latest system increased the professional development opportunities for teachers at the local and school level, but support to teachers' professional development on-the-job varies across schools, and support practices are often not leveraged because of inadequacies of school administrators' leadership (MECSS & JICA, 2018).

An in-depth look at the policies suggests that there is a need for strengthening the alignment of teacher policies and the enforcement of implementation.

As Mongolia continues to struggle to find better policies for better teachers, the government approved a teacher reform program called the “Skilled teacher” at the beginning of 2021. This is a measure to improve pre-service education; provide continuous development of teachers through support for local schools to build professional learning groups; and to increase teacher salaries.

Education systems try to support teachers with training or assessment; instead, they should enhance practice focusing on teaching and development (Bowe & Gore, 2017). Policies and mechanisms such as teacher training or curriculum documentation are important. However, what is more important is what is happening in the “black box” of the classroom to show impact of students’ learning and achievement. Reforms of the past focused on teachers rather than on teaching. Now, monitoring of teaching and lesson quality through lesson observations, and introduction of performance-based salary system might influence teachers’ practice and behavior in a way that might lead to an improvement in students’ learning and achievement.

References

- ADB. (2008). *Mongolia: Education sector reform project* (Technical Assistance Consultant’s Report).
- ADB. (2017). *Mongolia: Education Sector Development – Addressing school factors that affect access to quality education. Final reporty (Policy note)*.
- Bowe, J., & Gore, J. (2017). Reassembling teacher professional development: The case for quality teaching rounds. *Teachers and Teaching, 23*(2), 352–366.
- Chun, N., & Gentile, E. (2020). *Taking education to the next level: What can be learned from benchmarking education across economies?* ADB.
- Chun, S., Ulziisaikhan, G., Kim, D., Amarjargal, A., & Kim, J. (2020). A study on the applicability of ICALT for comparing teaching skills between secondary schools in Korea and Mongolia. *The SNU Journal of Education Research, 29*(1), 53–75.
- Enkhtuvshin, S. (2014). An evaluation of competency of initial teachers. In *Teacher competence in new century*. Mongolian National University of Education.
- Enkhtuvshin, S. (2020). Teacher education program efficiency and teaching preparedness. *Lavai, 24*, 110–144.
- Gore, J. M., Griffiths, T., & Ladwig, J. G. (2001). Productive pedagogy as a framework for teacher education: Towards better teaching. In *AARE annual conference*. Fremantle.
- Jadamba, B., Narantsetseg, D., Batdelger, J., & Baigalmaa, C. (2014). Studying teacher competencies in the new century from the universal approach. In *Teacher competence in new century*. Mongolian National University of Education.
- Kim, E., Doljinsuren, P., & Kim, G. (2017). The teacher quality management of teacher selection and qualification; an analysis of Mongolian case. *Korean Journal of Comparative Education, 27*(4), 231–254.
- Luvsandorj, T., & Oyun-Erdene, B. (2015). Measure of teacher competencies. In *Issues, solutions and alternatives of education measurement*. Mongolian National University of Education.
- MECSS and JICA. (2018). *Report on curriculum development, implementation and evaluation of primary and secondary education in Mongolia*. Ulaanbaatar.
- Pagma, B., Nergui, N., Monkhor, D., Jargalmaa, T., Hajidsuren, A., & Weidman, J. C. (2002). *Mongolia Country study. RETA No. 5946-REG: Sub-regional cooperation in managing education reforms*. Manila, Philippines.

- Steiner-Khamsi, G. (2005). Vouchers for teacher education (Non) reform in Mongolia: Transitional, postsocialist, or antisocialist explanations? *Comparative Education Review*, 49, 148–172.
- Steiner-Khamsi, G., & Stolpe, I. (2006). *Educational import: Local encounters with global forces in Mongolia*. Palgrave Macmillan.
- Ulziisaikhan, G. (2017). A study of teacher's satisfaction in Mongolia (in Mongolian). *Lavai*, 18–24.
- UNESCO. (2019). *Education in Mongolia: A country report*.
- van de Grift, W., Helms-Lorenz, M., & Maulana, R. (2014). Teaching skills of student teachers: Calibration of an evaluation instrument and its value in predicting student academic engagement. *Studies in Educational Evaluation*, 43, 150–159.
- World Bank. (2006). *Financing of education in Mongolia: Equity and efficiency implications*. World Bank.

Amarjargal Adiyasuren is a lecturer at Mongolian National University of Education. She formerly worked in Teachers' Professional Development Institute and Curriculum Reform Unit affiliated to Ministry of Education and Science. She worked in various national research projects related to school management, curriculum, pedagogy and assessment. She has been involved in comparative study of assessment of transversal skills with the Network on Education Quality Monitoring in the Asia-Pacific in the UNESCO Asia-Pacific and the Brookings Institution of the USA. She holds bachelor and master degree in Education from the University of Tokyo.

Ulziisaikhan Galindev is a senior lecturer in The Department of Educational Administration, Mongolian National University of Education. He received his master and doctoral degrees in Educational administration from Chungnam National University, South Korea. His current research interests and expertise cover education finance, education policy and teacher professional development.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 12

An Assessment of the Learning Environment and Teacher Interpersonal Behaviour at the Teacher Education Level



Adit Gupta  and Priya Sharma

Abstract The Indian teacher education scenario has undergone numerous changes in the last few years especially with the shift to the two-year teacher preparation programmes. As a result of this change, both the teacher educators and the student teachers had to adapt to the modified curriculum, teaching methodologies and assessment process. This paper focuses on assessing student teachers' perceptions about their classroom learning environments and teacher interpersonal behaviour. The study utilises the modified version of the What Is Happening In This Classroom (WIHIC) questionnaire and the Questionnaire on Teacher Interaction (QTI). The data was collected from 150 student teachers from a teacher education college studying in the third and fourth semester of the two-year B.Ed./B.Ed. Special Education programme. The results show that student teachers positively perceived their classroom learning environments. They expressed a lot of student cohesiveness, teacher support for the students, task orientation and involvement of students in the classroom activities. Students perceived an environment that promotes innovation, equity and a high level of cooperation. Results for teacher interpersonal behaviour show that student teachers perceived their teacher educators as good leaders who understand their needs. They are helpful and friendly and provided ample opportunities for students to express themselves freely. They also give responsibility to accomplish different tasks. The negative aspects of teacher interpersonal behaviour like uncertainty, admonishing and dissatisfied behaviour were given a low rating by the student teachers. They, however, felt that the teacher educators were strict in the class. Data analysis reveals that no significant associations exist between academic achievement and classroom learning environments and teacher interpersonal behaviour. Results also show that there were no significant gender differences in the learning environments. However, there were significant gender differences in the teacher interpersonal behaviour in favour of female student teachers. Also, no semester and programme based differences in the classroom learning environments and teacher interpersonal behaviour exist at the teacher education level.

A. Gupta (✉) · P. Sharma
MIER College of Education, Jammu, India
e-mail: adit@mier.in; priya.1905001@miercollege.in

Keywords Learning environment · Teacher interpersonal behaviour · Teacher education · WIHIC · QTI · Student teachers

1 Introduction

It is an undeniable fact that students spend a vast amount of time in class. As such, they have a stake in what happens to them in class and the perceptions of their experiences in the classrooms are of great significance. An evaluation of the educational process is not complete without the assessment of the learning environment in which students immerse themselves for many hours of their lives (Tan, 2011). On a daily basis in the classroom, teacher educators and student teachers assume different and complementary roles. The teacher educators' role is that to help their student teachers to reach educational and didactical objectives, while the student teachers' role is that to respond to the teacher educators' requests and to meet their learning expectations. The interactions between teacher educators and student teachers are regular and significant for reaching a common goal. However, every teacher educator displays a particular behaviour, that is different from that of his or her colleagues and that student teachers may or may not appreciate. And the reverse is also true: there are teacher educators who particularly like or dislike some student teachers' behavioural repertoires (Passini et al., 2015).

The teaching learning process cannot take place in a vacuum. In formal educational settings, it occurs as a result of interaction among members of the classroom. In classroom settings, elements of teaching-learning process include teacher educators, student teachers, content, learning process and learning situation. The learning situation or learning environment means the conditions in which learning take place (Malik & Rizvi, 2018). The teacher educator is considered a central figure in any classroom learning environment especially in Indian school/college settings, where the teacher educator controls the teaching-learning process and directs the activities of students on a day to day basis. Thus, the interaction which teacher educators have with their student teachers determines the nature of their interpersonal relationship and enables the teacher educator to improve their teaching practices. Getzels and Thelen (1960) suggested that teacher-student interaction is a powerful force that can play a major role in influencing the cognitive and affective development of students (Gupta & Fisher, 2011).

The teacher educator is the most important element in any educational program. It is the teacher educator who is mainly responsible for implementation of the educational process at any stage (NCTE, 1998). Teacher Education Curriculum is designed keeping in view the National Curriculum Framework of School Education. Reforms in teacher education focussed on the production of qualified and competent teachers at elementary stage (Classes 1–7) as well as secondary stage (Classes 8–10). Academic and professional standards of teacher education are to be ensured through development of well-planned teacher education programme with suitable implementation strategies. The curriculum framework of Teacher Education

developed by the NCTE, 2009, emphasized on provision of suitable curricular practices to the student teachers in various areas, such as:- understanding children and relating to them; understanding self and engagement self, engagement in critical reflection and innovation by student teachers; engagement with subject content and its linkage with learners' environment; development of professional skills in pedagogy and organization of various teaching learning activities inside and outside schools. The Curriculum Framework of NCTE covered three major areas, viz.: Foundations of Education, Curriculum and Pedagogy and School Internship. The NCTE Regulation, 2014, insisted on implementing National Curriculum framework of Teacher Education through longer duration of teacher education courses and accommodating various forms of integrated approaches in teacher preparation at elementary level and secondary level. The NCTE has prepared the guidelines for implementation of Integrated BA.B.Ed / B.Sc.B.Ed programme, Integrated B.Ed. and M.Ed. programme and other areas. Curricular areas have been expanded and duration of B.Ed. and M.Ed. programmes have been doubled from two semesters to four semesters (Sahoo & Sharma, 2018).

National Council for Teacher Education (NCTE) has overhauled the teacher education programmes at the graduation and post-graduation levels in India in the year 2014 in order to meet the needs of the twenty first century. The long awaited move of extending the duration of these courses from 1 year to 2 years was the most significant. An extension of the duration of the programmes resulted in complete restructuring of the teacher education programme. New courses like Language across Curriculum, courses for Enhancing Professional Competencies, Gender, School and Society and Creating an Inclusive School were welcome additions to the B.Ed. programme. Also, the period of school internship for B. Ed. students has been extended to 14 weeks which previously ranged from 5 to 6 weeks (Areekkuzhiyil, 2019). This modification resulted in an extended period of exposure for the student teachers to the actual school environment helping them acquire a comprehensive understanding of the functioning of schools. This recommendation is based on the assumption that longer duration programmes will provide sufficient time and opportunity for rigorous engagement of the future professionals—in view of a larger objective of professionalizing teacher education.

With the implementation of the two-year teacher education programme in the country, there have been numerous changes in the teaching-learning process, the way internships are conducted, major modifications in the curriculum and in the assessment and evaluation of pre-service teachers. This study aims to understand how these changes have impacted the classroom learning environments and the teacher-student interactions in the two-year teacher education programme especially in the courses of Bachelor of Education (B.Ed.) and the Bachelor of Education (B.Ed.) special Education programmes. Moreover, these courses have been converted to the semester system as compared to the annual system earlier, which has led to a lot more papers to be studied, sessional/projects to be submitted and examinations to be given. Hence, the study aims to assess if there are any course and semester-based differences in the perceptions of students regarding their learning environments and teacher-student interactions. It was also decided that since the

students study a plethora of papers/subjects in their course, the researchers would focus only on the core papers in education which are referred to as “Perspective Papers” while assessing the learning environments and teacher student interactions.

2 Review of Related Literature

2.1 *Research Studies on Classroom Learning Environment Using WIHIC*

Adnan et al. (2014) conducted a study on the learning environment and mathematics achievement of students at high performance schools (HPS). The purpose of this research was to determine the learning environment and mathematics achievements of High Performance Schools (HPS) students. A total of 362 Form Four students participated in the study. It was conducted using the survey methodology, with a set of questionnaires which was divided into Sections A and B. Section A consists of demographic-based questions to find out respondents’ background information. Section B is the What Is Happening In This Class (WIHIC) instrument which consists of 40 items to examine students’ perception of student closeness, teacher support, involvement, cooperation and fairness in the classroom learning environment. In addition, the students’ mathematics achievement was based on the grades of their final examination. The preliminary study produces a Cronbach’s Alpha of 0.952 for WIHIC. Data were processed using SPSS Windows Version 20.0 analysed to obtain percentage, frequency, mean, standard deviation, t-test and Pearson correlation. The data on students’ perception of their learning environment shows that the element of student closeness has the highest mean value, followed by the elements of cooperation, fairness, teacher support and involvement. In terms of students’ perception of the elements of involvement, cooperation and fairness, the results showed that there was a significant difference between male and female students. In addition, the study also found that there was a significant relationship between the elements of teacher support and fairness, and mathematics achievement.

Skordi (2014) conducted a study on “Learning Environment of University Business Studies Classrooms: Its Assessment, Determinants and Effects on Student Outcomes.” This study used the What Is Happening In this Class? (WIHIC) questionnaire, Revised Statistics Anxiety Rating Scale (RSARS) and Test of Statistics Related Attitudes (TOSRA) to assess perceptions of classroom environment, anxiety and attitudes among 375 students from 12 classes taking business statistics in Southern Californian universities. Students’ achievement also was measured by the final score for the course. When a three-way MANOVA revealed no interactions between three determinants (namely, sex, ethnicity and age) of student outcomes (anxiety, attitudes and achievement), sex, ethnic and age differences were interpreted independently. Relative to males, females had significantly higher scores for Task Orientation, Normality of Statisticians and the two anxiety scales. Relative to

younger students (22 years or less), older students perceived significantly more classroom Teacher Support and Involvement but had higher Learning Statistics Anxiety and lower achievement. Regarding statistically significant ethnic differences, Hispanics had lower achievement than Whites or Asians, and Asians perceived lower Task Orientation and Equity than Whites or Hispanics. Effect sizes for significant sex, ethnic and age differences typically ranged from approximately a quarter to a half a standard deviation (representing small to modest effects). Simple correlation and multiple regression analyses revealed statistically significant bivariate and multivariate associations between some of the WIHIC's learning environment scales and each of the student outcomes of statistics anxiety, attitudes and achievement. In particular, with other WIHIC scales mutually controlled, regression coefficients revealed that specific WIHIC scales were significant independent predictors of student outcomes.

Yang (2015) conducted a study on "Rural junior secondary school students' perceptions of classroom learning environments and their attitude and achievement in mathematics in West China". This paper reports findings from a survey of how rural junior secondary school students in the western part of China perceive their mathematics classroom learning environments and associations of learning environment with their attitudes toward mathematics and mathematics achievement. Using adaptations of the widely-used What Is Happening In this Class questionnaire and a mathematics attitude scale, the study involved data from 749 Grade 7, 842 Grade 8 and 864 Grade 9 students from 12 coeducational schools and 52 classrooms in three provinces. Data were analysed through factor analysis, descriptive statistics, two-way ANOVA, simple correlation analysis and multiple regression analysis. It was found that rural junior secondary students generally did not perceive their mathematics classroom environment very favourably, and they did not hold very positive attitudes towards mathematics. There existed significant gender and grade differences in the perceptions of mathematics classroom learning environments and attitudes towards mathematics. Positive correlation between mathematics classroom learning environment and students' attitudes towards mathematics and their mathematics achievement were identified.

Khalil and Aldridge (2019) conducted a study on Assessing students' perceptions of their learning environment in science classes in the United Arab Emirates. The sample included 784 students in 34 lower-secondary science classes in eight public schools in Abu Dhabi, UAE. The findings supported the validity of the dual-language Arabic/English version of the What Is Happening In this Class? (WIHIC) when used in this context. Also, all five learning environment scales were statistically significantly ($p < 0.01$) and positively related to each of eight attitudinal and engagement outcomes. This study has extended past research in the field of learning environments as the first of its kind to investigate the impact of cooperative learning in science classes on a range of student outcomes in the UAE. Methodologically, this study could be of significance to other researchers who might benefit from the availability of an Arabic version of the modified WIHIC for use in other studies.

2.2 *Research on Teacher-Student Interactions Using the QTI*

Gupta and Koul (2014) conducted the first study using QTI at the teacher education level for assessing teacher educators' interpersonal behaviour in a teacher education classroom setting in India. The Questionnaire on Teacher Interaction (QTI) was used with a sample of 270 students in an Indian teacher education college from the Jammu region (Jammu & Kashmir State, India) with respect to four compulsory papers being taught as part of the teacher education curriculum approved by the university. The results showed that the student teachers perceive their teacher educators' as good leaders most of the time and have also rated their teacher educators in terms of exhibiting a helpful and friendly nature, understanding and giving students a reasonable amount of freedom and responsibility in the classroom. The results also illustrate that the negative aspects of teacher-student interaction as assessed using QTI have been rated quite low by the student teachers as their teacher educators seldom exhibit admonishing behaviour, are less dissatisfied and less uncertain.

Fatima (2015) investigated one of the key elements of quality teaching, the teacher interpersonal behaviour and its impact on pre-service teachers' self-regulatory engagement. Data was collected with two extensively used instruments Questionnaire on teacher interaction QTI and Motivated strategies for learning questionnaire MSLQ. Data analysis revealed that only two of the dimensions have significant negative effect on self-regulatory engagement of student teachers.

Laudadío and Mazzitelli (2018) conducted a study on "Adaptation and validation of the Questionnaire on Teacher Interaction in Higher Education". This work aims at evaluating the validity and reliability of the Questionnaire on Teacher Interaction (QTI) applied in higher education by Soerjaningsih, Fraser and Aldridge. This instrument includes 48 items and enables the identification of the teacher's predominant behaviour according to two dimensions: proximity (cooperation-opposition) and influence (domination-submission). The questionnaire was applied to 256 students attending the first 2 years of courses of study related to Natural Science and Health Science at both public and private universities in the province of San Juan (Argentina). To evaluate the reliability, the Cronbach Alpha was applied, and the validity of the construct was studied by making a factorial analysis. The results indicate the existence of a two-dimensional structure: factor 1 is constituted by items that evaluate the proximity of the student-teacher relationship; it includes positive items that correspond to the cooperation sub dimension and negative items that correspond to the opposition sub dimension. Factor 2 is constituted by items that evaluate the influence in relation to domination. As regards reliability, when studying Factor 1, a Cronbach Alpha of .92 was obtained for cooperation and a Cronbach Alpha of .84 for opposition. Factor 2 had an Alpha of .61. The self-report globally shows an acceptable level of reliability. Summarising, favourable evidence was obtained about the discrimination of the items: factorial validity and the instrument's reliability. These results are important to understand the dynamics of the processes implied in the student-teacher relationship. Taking into account these results, it is considered that the QTI can be used as a guide to improve the

interpersonal relationships and to help teachers in their professional development. Using this instrument can be a valuable tool, both for investigation and intervention and prevention programs.

Ganapati et al. (2019) conducted a study on the teacher-student relationship and its impact on the behaviour of high school students. The purpose of this study was to investigate the teacher-student relationship and its impact on the behaviour of High school students. The objectives were to know the teachers' attitudes towards students and its impact to bring positive as well as negative behaviour change in the students. 50 high school students; 25 girls and 25 boys were taken and interview schedule is used. The study has reported that students often face emotional problems when negatively approached by the teachers. It is recommended that to create awareness among teachers in the school for the smooth handling the children with positive approaches.

Research studies in the area of classroom learning environments and teacher interpersonal behaviour are abundant both at the secondary as well as higher secondary level. However, in the field of learning environments and teacher student interactions there are very few research studies at the teacher education level. Due to lack of research work in this field, there is very little information regarding the quality of the teacher education programmes and how learning environments affect the student outcomes in teacher education classrooms especially in the Indian context. Thus, there is a need to study learning environments and teacher interpersonal behaviour at the B.Ed. level by assessing perceptions of student teachers of their classroom learning environment and also the interactions between teacher educators and their student teachers.

3 Objectives of the Study

The specific objectives of the study are:

1. To assess student-teachers' perceptions of their teacher education classroom learning environments.
2. To assess student-teachers' perceptions of their teacher educators interpersonal behaviour.
3. To investigate associations between classroom learning environments and academic achievement of student-teachers' in the perspective papers of the teacher education programme.
4. To investigate associations between teacher educators' interpersonal behaviour and academic achievement of student-teachers' in the perspective papers of the teacher education programme.
5. To investigate whether gender differences exist in classroom learning environments and teacher interpersonal behaviour in the teacher education programme.
6. To investigate whether semester differences exist in classroom learning environments and teacher interpersonal behaviour in the teacher education programme.

7. To investigate whether course differences exist in classroom learning environments and teacher interpersonal behaviour in the the teacher education programme.

4 Sample for the Study

In this study the researcher made an attempt to study the student-teacher's perceptions of their classroom learning environments and teacher interpersonal behaviour in relation to their academic achievement in Perspective papers at the B.Ed. level. For this purpose, a sample of 150 student-teachers' (both males and females) from a teacher education college of Jammu city were selected. The sample was chosen carefully so as to be representative of the population and comprised of both male and female student-teachers' in order to obtain an unbiased test of gender difference. Random Sampling technique was used in selecting the sample of the study.

5 Tools Used

After reviewing a number of instruments, the What Is Happening In This Classroom (WIHIC) (Fraser et al., 1996) and the Questionnaire on Teacher Interaction (QTI) (Wubbels et al., 1993) were selected to assess the classroom learning environments and teacher interpersonal behaviour at the B.Ed. level. The version of 'What Is Happening In This Class' (WIHIC) used in the study consists of 7 scales and 56 items (Fraser et al., 1996). The seven scales are Student Cohesiveness, Teacher Support, Involvement, Investigation, Task Orientation, Cooperation and Equity. The questionnaire was available in two forms, the Actual and the Preferred. The Actual Form measured the classroom environment in its current form while the Preferred Form measured perceptions of students' ideal or preferred classroom environments. The students responded to items using a five-point frequency response format (viz. Almost Never, Seldom, Sometimes, Often, Almost Always).

The WIHIC was modified for use with student teachers who were studying in a teacher education College. The only modification made to the WIHIC questionnaire was the removal of the Investigation Scale. The Investigation Scale in the WIHIC was primarily added to assess the perceptions of students in science/mathematics classrooms and did not serve any meaningful purpose in the assessment of teacher education programme. Finally, a new scale, namely, 'Innovation' which was taken from College and University Classroom Environment Inventory (Fraser et al., 1986) was added to assess the extent to which the instructor plans new, unusual class activities, teaching techniques, and assignments in the class. The investigators felt that including the Innovation Scale added value to the overall study as at the teacher education level, the teacher educators are using innovative methods of teaching and are employing information and communication technologies for teaching-learning

and assessment, that it would be apt to assess the perceptions of students towards these innovations in the classroom. The Innovation scale also consisted of eight items to which students responded using a five-point scale, i.e., the items were scored 1, 2, 3, 4, 5, respectively, for the Almost Never, Seldom, Sometimes, Often and Almost Always responses. The different scales of the modified version of the WIHIC are shown in Table 12.1.

The QTI enables information concerning student's perceptions of teacher interpersonal behaviour to be gathered. The original version of the QTI that was developed in the early 1980s in the Netherlands had 77-items (Wubbels et al., 1985). The Australian version developed by Wubbels et al. (1993) was used in this study. This 48-item short form of the QTI has six items for every sector of the model for teacher interpersonal behaviour. Each of the eight sectors describes a particular behaviour type. Responses to the items are scored 1, 2, 3, 4, 5, respectively, for the responses, Never, Seldom, Sometimes, Often and Always. The different scales of the QTI are shown in Table 12.2.

This version of QTI was used with school students so far. But in this study the QTI was used with student-teachers' studying in a teacher education college in Jammu. Therefore, there was a need to modify the items in the questionnaire to be used at the B.Ed. level so that the items were properly understood by the student-teachers' and they were able to respond in the right manner for e.g., item number 3 of the Uncertain scale read, 'This teacher seems uncertain' which was changed to 'This teacher seems uncertain about students' activities in the class'. Similarly, item number 9 of the Leadership scale read, 'This teacher holds our attention' which was changed to 'This teacher holds our attention in the class'.

Both the tools were modified for use in the present study; hence, their reliability and validity were established. The Cronbach alpha reliability coefficient was utilised as a scale internal consistency metric, demonstrating how consistent the test

Table 12.1 Names and descriptions of modified WIHIC scales

Scale name	Scale description
Student cohesiveness (SC)	The extent to which student know, help and are supportive of one another.
Teacher support (TS)	The extent to which the teacher helps, befriends, trusts and is interested in students.
Involvement (IV)	The extent to which students are attentive interest, participate in discussions, do additional work and enjoy the class.
Task orientation (TO)	The extent to which it is important to complete the activities planned and stay on the subject matter.
Innovation (INN)	The extent to which the instructor plans new, unusual class activities, teaching techniques, and assignments.
Cooperation (CO)	The extent to which students cooperate rather than compete with one another on learning tasks.
Equity (EQ)	The extent to which students are treated equally by the teacher.

Responses of the items are scored 1, 2,3,4,5 respectively, from almost never, seldom, sometimes, often to almost always. Missing or invalid responses are scored 3, the mid-range value

Table 12.2 Description of Items for Each Scale in the QTI

Scale	Description
Leadership	Extent to which teacher provides leadership to class and holds student attention.
Helping/friendly	Extent to which the teacher is friendly and helpful towards students.
Understanding	Extent to which teacher shows understanding and care to students.
Student responsibility/freedom	Extent to which the students are given opportunities to assume responsibilities for their own activities.
Uncertain	Extent to which teacher exhibits her/his uncertainty.
Dissatisfied	Extent to which teacher shows unhappiness/dissatisfaction with the students.
Admonishing	Extent to which the teacher shows anger/temper and is impatient in class
Strict	Extent to which the teacher is strict with demands of the students.

items are when compared to other test items that assess the same construct of interest. A discriminant validity index (the mean correlation of a scale with other scales) was utilized to show that each WIHIC scale estimates a different aspect from the other scales in the questionnaire. For the WIHIC scale the reliability values ranged from 0.75 for the Innovation Scale to 0.88 for the Involvement scale for the actual form of the questionnaire. For the preferred form of the questionnaire the reliability coefficient values ranged from 0.75 for the Student Cohesiveness and Innovation scale to 0.89 for the Task Orientation and Cooperation scale. For the Questionnaire on Teacher Interaction (QTI), the reliability coefficient values ranged from 0.66 for the Understanding Scale to 0.83 for Leadership and Uncertain scale. The WIHIC's and QTI's reliability values were consistently above 0.50. This suggested that the WIHIC and QTI can be regarded as a reliable tool (De Vellis, 1991) with teacher trainees in the B.Ed. and B.Ed. Special Education courses. Similarly, in the Actual Form, the discriminant validity results for the seven WIHIC scales ranged from 0.49 for the Teacher Support and Innovation scales to 0.54 for the Task Orientation and Equity scales, and between 0.45 for the Teacher Support scale to 0.56 for the Task Orientation, Cooperation, and Equity scales in the preferred form. In general, the results of reliability and validity corroborate the circumplex model of the QTI and hence validate it for use at the teacher education level.

Apart from the above mentioned tools, the researchers also collected data on the achievement of student teachers in terms of their performance in the end-semester examinations in the perspective papers being studied by them. The marks obtained by the student teachers were used for purpose of investigating the associations between the students marks and their classroom learning environments and teacher student interactions.

6 Results of the Study

6.1 Means and Standard Deviations of the WIHIC

To answer Research Question 1 “To assess student-teachers’ perceptions of their teacher education classroom learning environments”, the data on the seven scales of the What is Happening in this Class (WIHIC) questionnaire were collected from 150 student-teachers’ who have been studying in a B.Ed. College. Item means and standard deviations were computed to determine the nature of classroom learning environment using the WIHIC. The statistical significance of the difference between means (t-test) was also calculated to study whether the differences in the means of the Actual and Preferred Forms of the WIHIC when used in a teacher education classroom setting were significant. The data obtained are presented in Table 12.3. The results show that the mean scores of the different scales of the WIHIC ranged from 3.73 for the Teacher-support scale to 4.23 for the Task Orientation scale in the Actual Form which shows that student-teachers’ were generally able to complete their classroom activities in a planned manner and were also able to stay on their subject matter in the teacher education classroom. The mean scores of Student Cohesiveness scale is 4.15, Involvement scale is 3.75, Innovation scale is 3.75, Cooperation scale is 4.21 and Equity scale is 4.06 which indicates that the student-teachers’ know each other very well and are supportive of one another, they remain attentive in the class and give their opinions during class discussions, new teaching techniques and activities are planned by their teacher educators in the class, they cooperate with one another while doing assignments and class activities and every student-teacher gets the same opportunity to contribute to class discussions.

An examination of the mean scores in the Preferred Form of the WIHIC shows that the value ranged from 3.71 for the Teacher-support scale to 4.25 for the Task Orientation scale. This indicates that student-teachers’ usually want to complete their activities in a planned manner and also want to stay on the subject matter in the teacher education classroom. The values of the standard deviations in both the

Table 12.3 Means, Standard Deviations (SD) and Significance of Difference between Means (t) for the WIHIC

Scale name	No. of Items	Mean		Standard deviation(SD)		t
		Act.	Pref.	Act.	Pref.	
Student cohesiveness	8	4.15	4.08	0.62	0.64	1.56
Teacher support	8	3.73	3.71	0.78	0.76	0.51
Involvement	8	3.75	3.78	0.80	0.75	0.77
Innovation	8	3.75	3.76	0.70	0.70	0.03
Task orientation	8	4.23	4.25	0.71	0.75	0.48
Cooperation	8	4.21	4.24	0.69	0.74	0.84
Equity	8	4.06	4.17	0.81	0.74	2.89**

N = 150 **Significant at 0.01 level

Actual and Preferred Form of the WIHIC are less than 1, which suggests that there are no major deviations in student-teachers' perceptions of their classroom learning environment.

The results for the paired t-tests indicated that there is a significant difference ($p < 0.001$, $p < 0.01$, $p < 0.05$) between the actual and preferred means for only one scale out of the seven scales of WIHIC, i.e., Equity with a t value of 2.89. Thus, there is a significant difference between the actual and preferred means for the scale which shows that student-teachers' want more attention and equal treatment from the teacher educator in the classroom.

7 Means and Standard Deviations of the QTI

To answer the Research Question 2, the data for the descriptive statistics concerning the Questionnaire on Teacher Interaction (QTI) were collected from 150 student-teachers' studying in a B.Ed. college and the values of means and standard deviations are given in Table 12.4. The highest mean value is 4.23 for the Leadership scale and the least value is 2.54 for the Admonishing scale.

The overall analysis of the results in Table 12.4 shows that the student-teachers' see their teacher educators as good leaders most of the time and have also rated their teacher educators in terms of exhibiting helpful and friendly nature, understanding and giving students freedom and responsibility in the classroom. In fact, the positive factors have been exhibited by the teacher educators quite often in the classroom. One interesting feature of the analysis is that student-teachers' perceive their teacher educators to be strict which is acceptable in India as the teacher educator is in charge of a class and gives direction to the student teachers in various academic matters. Also, the negative aspects of the teacher-student interaction have been rated quite low by the student educators as teacher educators seldom exhibit admonishing behaviour, are less dissatisfied and less uncertain. Figure 12.1 represents a sector profile depicting student's perception of the teacher-student interpersonal behaviour at the B.Ed. level which was developed by plotting the mean scores of the eight

Table 12.4 Means and standard deviations for the QTI

Scale name	No. of items	Mean	S.D
Leadership (DC)	6	4.23	0.66
Understanding (CS)	6	4.13	0.60
Uncertain (SO)	6	2.57	0.97
Admonishing (OD)	6	2.54	0.88
Helping/friendly (CD)	6	3.87	0.63
Student responsibility/freedom (SC)	6	3.28	0.75
Dissatisfied (OS)	6	2.58	0.94
Strict (DO)	6	3.19	0.77

N = 150

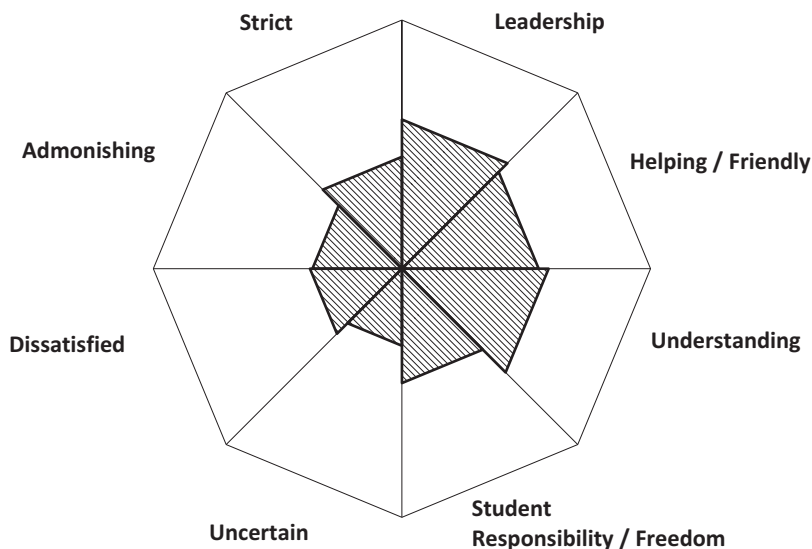


Fig. 12.1 Sector profile diagram of student-teachers' perception of their teacher educators' interpersonal behaviour

scales of the QTI (student questionnaire) in an excel worksheet. The sector profile reveals diagrammatically the degree to which students perceive each behavioural aspect exhibited by the teacher educator as measured through the QTI.

From Table 12.4 we can see that the standard deviation ranges from 0.60 for the Understanding scale to 0.97 for the Uncertain scale. Since the values of the standard deviation are less than 1.00, it suggests that there is no major diversity in students' perceptions.

8 Associations with the WIHIC

8.1 Association of Students' Perception of Their Classroom Learning Environment with Academic Achievement

The association between the academic achievement of the student-teachers' and the perceptions of their classroom learning environments as measured by the WIHIC were also explored using simple and multiple correlations followed by the computation of the regression coefficient. The statistical results to answer Research Question 3 are presented in Table 12.6.

The data for academic achievement was taken from the semester end result of the student-teachers' in the perspective papers of the B.Ed. programme. The data illustrated in Table 12.5 indicates that for simple correlation (r) all the seven scales of

Table 12.5 Associations between the WIHIC Scales and Academic Achievement in terms of Simple Correlation (*r*), Multiple Correlation (*R*) and Standardised Regression Coefficient (β)

Scale Name	Academic Achievement	
	<i>r</i>	β
Student cohesiveness	0.13	0.14
Teacher support	0.07	0.23
Involvement	-0.08	-0.28
Innovation	-0.07	-0.21
Task orientation	0.07	-0.03
Cooperation	0.13	0.23
Equity	0.06	0.01

Multiple Correlation $R = 0.32^*$ $R^2 = 0.10$

* Significant at $p < 0.01$, $N = 150$

Table 12.6 Associations between the QTI Scales and the academic achievement in terms of Simple Correlation (*r*), Multiple Correlation (*R*) and Standardised Regression Coefficient (β)

Scale name	Academic achievement	
	<i>r</i>	β
Leadership	-0.02	-0.04
Understanding	0.01	-0.02
Uncertain	0.02	-0.08
Admonishing	0.08	0.29
Helping/friendly	0.04	0.15
Student responsibility/freedom	0.04	-0.03
Dissatisfied	-0.02	-0.08
Strict	-0.07	-0.17

Multiple Correlation $R = 0.19$ $R^2 = 0.04$
 $N = 150$

the WIHIC are not statistically significant and are not positively associated with student-teachers’ academic achievement ($p < 0.001$, $p < 0.05$, $p < 0.01$) at the individual level of analysis. The values of correlation ranged from 0.06 for the Equity scale to 0.13 for the Cooperation scale. Thus, academic achievement is not significantly correlated in a positive direction with any of the seven scales, which implies that there is no positive relationship between classroom learning environment and academic achievement of the student-teachers’ in terms of their performance in the examination and attainment of knowledge.

The multiple correlation (*R*) between student-teachers’ perceptions as measured by the different scales of WIHIC and the Academic Achievement scale (as seen in Table 12.5) is 0.32 at the individual level of analysis, which is statistically significant ($p < 0.01$). The R^2 value indicates that 10 percent of the variance in the student-teachers’ academic achievement can be attributed to the classroom learning environment. Standardized regression values were calculated to provide information about the unique contribution of each learning environment scale to the Academic Achievement scale. Regression coefficient values (β) (as given in Table 12.5) indicate that none of the seven WIHIC scales uniquely account for a

significant ($p < 0.001$, $p < 0.01$, $p < 0.05$) amount of variance in academic achievement. It is evident from the data that the classroom learning environment at the B.Ed. level may not help in improving the academic achievement of the student-teachers' as both the correlation and regression coefficients do not have a positive and significant association with the academic achievement scores.

9 Associations with the QTI

9.1 Association of Students' Perception of their Teacher-Student Interactions with Academic Achievement

Simple (r) and multiple correlation (R) along with computation of the regression coefficient (β) were used to study the associations between the student-teachers' perceptions of their teacher educators interpersonal behaviour as measured by the QTI and their academic achievement. Table 12.6 illustrates the results of the statistical computation for Research Question 4.

Analysis of data shows that none of the eight scales of the QTI have a significant correlation with the academic achievement scores. The correlation values for the scales of QTI range from -0.01 for the Understanding scale to 0.08 for the Admonishing scale. The multiple correlation (R) between student-teachers' perceptions as measured by the different scales of the QTI and the academic achievement scores (as seen in Table 12.6) is 0.19 at the individual level of analysis, which is statistically not significant. The R^2 value indicates that just 4% of the variance in the academic achievement can be attributed to the teacher educator's interpersonal behaviour. Standardized regression values were calculated to provide information about the unique contribution of each QTI scale to the academic achievement scores. Regression coefficient values (β) indicate (see Table 12.6) that none of the eight QTI scales uniquely account for a significant ($p < 0.001$, $p < 0.01$, $p < 0.05$) amount of variance in academic achievement scores. It is evident from the data that the teacher-student interactions at the B.Ed. level may not help in improving the academic achievement of the student-teachers'.

10 Gender Differences

The fifth research question was to investigate whether gender differences exist in classroom learning environments and teacher-student interactions at the teacher education level. In the present sample of 150 students taken from the B.Ed. College, there were 144 (96%) female student-teachers' and 06 (4%) male student-teachers'. In this section, the gender differences with respect to classroom learning environments and teacher-student interactions have been discussed.

10.1 Gender Differences and Classroom Learning Environment

The means and standard deviations for each of the male and female groups were computed followed by a test of significance of difference between means (*t*-test for independent samples) on the seven scales of the WIHIC (research Question 5). The data obtained are presented in Table 12.7.

From the information given in Table 12.7, it can be seen that none of the seven scales of the WIHIC are statistically significant ($p < 0.001$, $p < 0.01$, $p < 0.05$). The *t* value for the WIHIC scales ranged from 0.02 for Involvement scale to 1.57 for Innovation scale. This means that no gender differences exist in classroom learning environments at the B.Ed. Level. Thus, both male and female students perceived their classroom learning environments in a similar manner, thus signifying homogeneity in the group. This also may be due to the fact that the sample of males and female was skewed in favour of the female student teachers as more female student teachers pursued the teacher education programme in Jammu.

10.2 Gender Differences and Perceptions of Teacher-Student Interaction

The means and standard deviations for the two groups were computed followed by a test of significance of difference between means (*t*-test for separate samples) to find out if there were any gender differences on the eight scales of the QTI. The data obtained statistically are presented in Table 12.8.

Table 12.7 Means, standard deviations and significance of difference between means for gender differences in students' perceptions of learning environment as measured by the WIHIC

Scale	Gender	Mean	Std. error mean	Standard deviation	<i>t</i>
Student cohesiveness	Females	4.15	0.05	0.62	0.17
	Males	4.10	0.30	0.74	
Teacher support	Females	3.73	0.06	0.78	0.13
	Males	3.77	0.33	0.82	
Involvement	Females	3.76	0.07	0.81	0.02
	Males	3.75	0.18	0.45	
Innovation	Females	3.77	0.06	0.71	1.57
	Males	3.31	0.27	0.66	
Task orientation	Females	4.23	0.06	0.72	0.28
	Males	4.31	0.17	0.42	
Cooperation	Females	4.20	0.58	0.69	0.81
	Males	4.44	0.23	0.57	
Equity	Females	4.05	0.07	0.82	1.04
	Males	4.39	0.15	0.37	

Females: N = 144; Males: N = 06

Table 12.8 Means, standard deviations and significance of difference between means for gender differences in students' perceptions of teacher-student interaction as measured by the QTI scale

Scale	Gender	Mean	Std. error mean	Standard deviation	<i>t</i>
Leadership	Females	4.22	.055	0.66	0.79
	Males	4.44	0.18	0.44	
Understanding	Females	4.12	0.05	0.61	0.95
	Males	4.36	0.16	0.38	
Uncertain	Females	2.59	0.08	0.98	1.19
	Males	2.11	0.27	0.67	
Admonishing	Females	2.55	0.07	0.89	0.76
	Males	2.27	0.22	0.53	
Helping/friendly	Females	3.86	0.05	0.64	0.62
	Males	4.02	0.14	0.34	
Student responsibility /freedom	Females	3.28	0.06	0.75	0.12
	Males	3.25	0.16	0.40	
Dissatisfied	Females	2.62	0.08	0.94	2.33*
	Males	1.72	0.19	0.48	
Strict	Females	3.22	0.06	0.76	2.67**
	Males	2.38	0.23	0.56	

Females: $n = 144$; Males: $n = 06$

From the information given in Table 12.8, it can be seen that out of the eight scales of the QTI only two scales, i.e., Dissatisfied with a t value of 2.33 and Strict with a t value of 2.67 are statistically significant ($p < 0.01$, $p < 0.05$). In these scales, females have a higher mean score than males. This means that female student-teachers 'seems dissatisfied and also find their teacher to be strict at the B.Ed. level as compared to male student-teachers'. This could be attributed to the fact that majority of the students are females and hence they have more interaction in the classroom as compared to the male students. Figure 12.2 represents the mean scores of the male and female students on the eight scales of the QTI.

11 Semester Differences

The sixth research question was to investigate whether semester differences exist in classroom learning environments and teacher-student interactions in the perspective papers at the B.Ed. Level. In the present sample of 150 students taken from the B.Ed. College, there were 72 (48%) student-teachers 'studying in Semester 3 and 78 (52%) student-teachers' studying in Semester 4. In this section, the semester differences with respect to classroom learning environments and teacher-student interaction have been discussed.

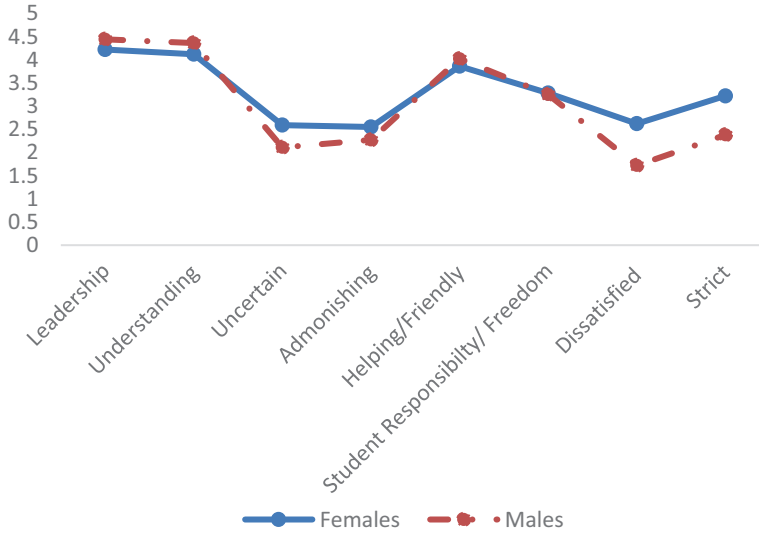


Fig. 12.2 Mean scores of male and female students on the eight scales of the QTI

11.1 Semester Differences and Classroom Learning Environment

The means and standard deviations for each of the semester 3 and semester 4 student-teachers’ were computed followed by a test of significance of difference between means (*t*-test for independent samples) on the seven scales of the WIHIC. The data obtained are presented in Table 12.9, which shows that none of the seven scales of the WIHIC are statistically significant ($p < 0.01$, $p < 0.05$, $p < 0.001$). The *t* value for the WIHIC scales ranged from 0.24 for Student Cohesiveness scale to 1.74 for Innovation scale. This means that no semester differences exist in classroom learning environments at the B.Ed. Level. Thus, student-teachers’ of both semesters perceived their classroom learning environments in a similar manner, thus signifying homogeneity in the group.

11.2 Semester Differences and Teacher-Student Interactions

The means and standard deviations for the two semesters were computed followed by a test of significance of difference between means (*t*-test for separate samples). The data obtained statistically are presented in Table 12.10. The data analysis reveals that there are no semester differences in student-teachers’ perceptions of their teacher-student interactions at the B.Ed. Level. Thus, student-teachers’ of both semesters perceived their teacher-student interactions in a similar manner.

Table 12.9 Means, standard deviations and significance of difference between means for semester differences in students' perceptions of learning environment as measured by the WIHIC

Scale	Semester	Mean	Std. error mean	Standard deviation	<i>t</i>
Student cohesiveness	Sem 3	4.13	0.07	0.63	0.24
	Sem 4	4.16	0.07	0.62	
Teacher support	Sem 3	3.68	0.09	0.81	0.64
	Sem 4	3.77	0.08	0.76	
Involvement	Sem 3	3.66	0.09	0.85	1.41
	Sem 4	3.84	0.08	0.75	
Innovation	Sem 3	3.65	0.08	0.69	1.74
	Sem 4	3.85	0.08	0.72	
Task orientation	Sem 3	4.19	0.08	0.75	0.67
	Sem 4	4.27	0.07	0.68	
Cooperation	Sem 3	4.18	0.08	0.69	0.54
	Sem 4	4.24	0.07	0.70	
Equity	Sem 3	4.01	0.09	0.81	0.76
	Sem 4	4.11	0.09	0.80	

Semester 3: N = 72; Semester 4: N = 78

Table 12.10 Means, standard deviations and significance of difference between means for semester differences in students' perceptions of teacher-student interaction as measured by the QTI Scale

Scale	Semester	Mean	Std. error mean	Standard deviation	<i>t</i>
Leadership	Sem 3	4.21	0.08	0.68	0.46
	Sem 4	4.25	0.07	0.64	
Understanding	Sem 3	4.04	0.07	0.62	1.83
	Sem 4	4.22	0.06	0.57	
Uncertain	Sem 3	2.51	0.11	1.01	0.79
	Sem 4	2.64	0.11	0.96	
Admonishing	Sem 3	2.47	0.10	0.87	0.96
	Sem 4	2.61	0.10	0.89	
Helping/friendly	Sem 3	3.91	0.07	0.62	0.78
	Sem 4	3.83	0.07	0.64	
Student responsibility /freedom	Sem 3	3.24	0.08	0.75	0.61
	Sem 4	3.32	0.08	0.75	
Dissatisfied	Sem 3	2.54	0.11	0.97	0.57
	Sem 4	2.63	0.10	0.92	
Strict	Sem 3	3.12	0.09	0.76	1.15
	Sem 4	3.26	0.09	0.77	

Sem 3: N = 72; Sem 4: N = 78.

12 Course Differences

The last research question was to investigate whether course differences exist in classroom learning environments and teacher-student interactions at the B.Ed. Level. In the present sample of 150 students taken from a B.Ed. College, there were 127 (84.6%) student-teachers' from the B.Ed. course and 23 (15.4%) student-teachers' from the B.Ed. Special Education course.

12.1 Course Differences and Classroom Learning Environment

The means and standard deviations for each of the B.Ed. and B.Ed. Special Education student-teachers' were computed followed by a test of significance of difference between means (*t*-test for independent samples) on the seven scales of the WIHIC. The data obtained are presented in Table 12.11. From the information given in Table 12.11, it can be seen that none of the seven scales of the WIHIC are statistically significant ($p < 0.001$, $p < 0.01$, $p < 0.05$). The *t* value for the WIHIC scales ranged from 0.14 for the Student Cohesiveness scale to 1.59 for the Equity scale. This means that no course differences exist in classroom learning environments at the B.Ed. Level. Thus, student-teachers' of both courses perceived their classroom learning environments in a similar manner.

Table 12.11 Means, Standard Deviations and Significance of Difference between Means for Course Differences in Students' Perceptions of Learning Environment as measured by the WIHIC

Scale	Course	Mean	Std. error mean	Standard deviation	<i>t</i>
Student cohesiveness	B.Ed.	4.14	0.05	0.63	0.14
	B.Ed. Spl	4.16	0.13	0.61	
Teacher support	B.Ed.	3.69	0.06	0.78	1.33
	B.Ed. Spl	3.92	0.15	0.73	
Involvement	B.Ed.	3.74	0.07	0.80	0.44
	B.Ed. Spl	3.82	0.16	0.79	
Innovation	B.Ed.	3.74	0.06	0.71	0.55
	B.Ed. Spl	3.83	0.14	0.70	
Task orientation	B.Ed.	4.21	0.06	0.72	0.77
	B.Ed. Spl	4.33	0.13	0.63	
Cooperation	B.Ed.	4.18	0.06	0.69	1.13
	B.Ed. Spl	4.36	0.14	0.69	
Equity	B.Ed.	4.01	0.07	0.83	1.59
	B.Ed. Spl	4.30	0.13	0.64	

B.Ed.: N = 127; B.Ed. Special Education: N = 23

12.2 Course Differences and Perceptions of Teacher-Student Interaction

The means and standard deviations for the B.Ed. and B.Ed. Special Education student-teachers' were computed followed by a test of significance of difference between means (*t*-test for separate samples) to find out if course differences exist in teacher-student interactions. The data obtained statistically are presented in Table 12.12.

The data analysis reveals that there are no course differences in student-teachers' perceptions of their teacher-student interactions at the B.Ed. Level. Thus, student-teachers' of both B.Ed. and B.Ed. Special Education perceived their teacher-student interactions in a similar manner, thus signifying homogeneity in the group.

13 Limitations of the Study

The main objective of this research was to assess student-teachers' perceptions of their classroom learning environments and teacher-interpersonal behaviour in relation to their academic achievement in perspective papers at the B.Ed. level. One of the limitations of this study was that the sample size was reduced as only one college was involved and the number of teacher educators was less, otherwise this study could have provided more information on the extent of teacher-student

Table 12.12 Means, standard deviations and significance of difference between means for course differences in students' perceptions of teacher-student interaction as measured by the QTI Scale

Scale	Course	Mean	Std. error mean	Standard deviation	<i>t</i>
Leadership	B.Ed.	4.21	0.05	0.66	0.84
	B.Ed. Spl	4.34	0.14	0.67	
Understanding	B.Ed.	4.09	0.05	0.61	1.84
	B.Ed. Spl	4.34	0.11	0.56	
Uncertain	B.Ed.	2.60	0.08	0.98	0.72
	B.Ed. Spl	2.44	0.20	0.97	
Admonishing	B.Ed.	2.54	0.07	0.89	0.24
	B.Ed. Spl	2.58	0.17	0.85	
Helping/friendly	B.Ed.	3.84	0.05	0.64	1.42
	B.Ed. Spl	4.04	0.10	0.52	
Student responsibility /freedom	B.Ed.	3.25	0.06	0.75	1.31
	B.Ed. Spl	3.47	0.15	0.72	
Dissatisfied	B.Ed.	2.59	0.08	0.95	0.20
	B.Ed. Spl	2.55	0.18	0.90	
Strict	B.Ed.	3.17	0.06	0.78	0.93
	B.Ed. Spl	3.33	0.14	0.67	

B.Ed.: N = 127; B.Ed. Spl: N = 23

interactions in multiple colleges. Although the statistical analysis of the questionnaire suggested that the WIHIC and QTI were valid tools for use in Indian classrooms it was felt that there was a need to modify the items in the QTI questionnaire as some items were not properly understood by the student-teachers'. This was overcome to some extent because after preliminary administration of the questionnaires, efforts were made to correct and improve those questions to which the student-teachers' did not respond well. In addition, due to COVID-19 pandemic, this study was confined to only 150 student-teachers' and a larger sample could have added to the richness of results. Another limitation of the study was that the achievement of the students was measured only on the basis of their marks obtained in the end-semester examination and did not cover a broad spectrum of activities based on non-academic activities.

14 Discussion and Conclusions

The results of the present study in the context of research in the field of classroom learning environment and teacher interpersonal behaviour in a teacher education college of Jammu city are considerable mainly because it is one of the few studies to use the What Is Happening In This Classroom (WIHIC) and Questionnaire on Teacher Interaction (QTI) at the B.Ed. level in India. The results have shown that positive classroom learning environment and teacher interpersonal behaviour exists in the teacher education classroom settings. The results of the study showed that student teachers perceived their classroom learning environments in a positive manner and expressed that there is lot of student cohesiveness, teacher support for the students, task orientation and involvement of students in the classroom activities. Data also shows that students perceived an environment that promotes innovation in the classroom, equity in treatment of students and high level of cooperation amongst students. Results for Teacher Interpersonal behaviour shows that student teachers perceived their teacher educators as good leaders, having understanding of students teacher' needs, helpful and friendly and provided ample opportunities for students to express themselves freely and also give responsibilities to accomplish different tasks. The negative aspects of teacher interpersonal behaviour such as uncertain, admonishing and dissatisfied behaviour was rated quite low by the student teachers. They however felt that the teacher educators were strict in the class. Data analysis further reveals that there were no significant associations between the academic achievement of the students and their classroom learning environments and teacher interpersonal behaviour. Also, no significant semester differences and programme-based differences have been reported in the classroom learning environments and teacher interpersonal behaviour at the teacher education level. The study is significant because the outcomes can provide guidelines for teacher educators to improve their classroom learning environments and teacher interpersonal behaviour at the

B.Ed. level. However, the findings of this study in terms of teacher interpersonal behaviour provides valuable feedback for the teacher educators to look at how they can modify their behaviour towards student-teachers' in the teacher education college and the areas they need to work on to make the classroom learning environment more effective. In a nutshell, the result from this study can provide guidelines for teacher educators who wish to develop more positive and productive classroom learning environments. The teacher educators will be able to use the results of the study to assess their own classroom learning environments and teacher-student interactions. This will help them in understanding those psychosocial aspects of their classroom which require improvement such as Student Cohesiveness, Teacher Support, Involvement, Innovation, Task Orientation, Cooperation and Equity. The assessment of the positive and negative aspects of teacher-student interactions shall also help the teacher educators in bring meaningful changes in their classroom transactions and behaviour that support constructive learning environments.

References

- Adnan, M., Abdullah, M. F. N. L., Puteh, M., Ahmad, C. N. C., & Maat, S. M. (2014). The learning environment and mathematics achievement of students of high performance schools (HPS). *Jurnal Pendidikan Matematik*, 2(1), 1–15.
- Areekkuzhiyil, S. (2019). Structural changes in teacher education Programmes: Some ground realities. *Online Submission*, 57(1), 66–68.
- De Vellis, R. F. (1991). *Scale development: Theory and application*. Sage Publications.
- Fatima, S. (2015). Effect of teacher interaction on self-regulatory engagement: Pakistani trainee-teachers' perspective. *Educational Research International*, 4(1), 87–108.
- Fraser, B. J., Fisher, D. L., & McRobbie, C. J. (1996). *Development, validation and use of personal and class forms of a new classroom environment instrument*. Paper presented at the annual meeting of the American Educational Research Association.
- Fraser, B. J., Tregust, D. F., & Dennis, N. C. (1986). Development of an instrument for assessing classroom psychosocial environment at universities and colleges. *Studies in Higher Education*, 11(1), 43–54.
- Ganapati, G., Laveena, G., & Mello, D. (2019). A study on the teach-student relationship and its impact on the behaviour of high school students. *International Journal of Case Studies in Business*, 3(1), 29–34. <https://doi.org/10.5281/zenodo.2589822>
- Getzels, J. W., & Thelen, H. A. (1960). A conceptual framework for the study of the classroom group as a social system. In A. Morrison and D. McIntyre (Eds.), *The social psychology of teaching* (pp. 17–34). Penguin.
- Gupta, A., & Fisher, D. (2011). Teacher-student interactions in a technology-supported science classroom environment in relation to selected learning outcomes: An Indian study. *MIER Journal of Educational Studies Trends & Practices*, 1(1), 41–59. <https://doi.org/10.52634/mier/2011/v1/i1/1625>
- Gupta, A., & Koul, R. (2014). Assessing teacher educators' interpersonal behaviour in a teacher education classroom setting in India. *Journal of Education and Pedagogy*, 6(1), 9–24.
- Malik, R. H., & Rizvi, A. A. (2018). Effect of classroom learning environment on students' academic achievement in mathematics at secondary level. *Bulletin of Education and Research*, 40(2), 207–218.

- National Council for Teacher Education. (1998). *NCTE manual on teacher education and teacher training*. NCTE.
- Khalil, N., & Aldridge, J. (2019). Assessing students' perceptions of their learning environment in science classes in The United Arab Emirates. *Learning Environments Research*, 22(3), 365–386.
- Laudadío, J., & Mazzitelli, C. (2018). Adaptation and validation of the questionnaire on teacher interaction in higher education. *Interdisciplinary: Journal of Psychology and Related Sciences*, 35(1). <https://doi.org/10.16888/interd.2018.35.1.8>
- Passini, S., Molinari, L., & Speltini, G. (2015). A validation of the questionnaire on teacher interaction in Italian secondary school students: The effect of positive relations on motivation and academic achievement. *Social Psychology of Education*, 18(3), 547–559. <https://doi.org/10.1007/s11218-015-9300-3>
- Sahoo, P. K., & Sharma, P. (2018). Student teachers' perception towards curriculum reform in teacher education programme in Odisha. *Educational Quest*, 9(1), 1–11. <https://doi.org/10.30954/2230-7311.2018.04.01>
- Skordi, P. G. (2014). *Learning environment of university business studies classrooms: Its assessment, determinants and effects on student outcomes* (Unpublished Doctoral Thesis). Curtin University.
- Tan, K. C. E. (2011). *Students' learning environment and attitudes toward science in light of the teach less, learn more initiative*. (Doctoral dissertation, Curtin University of Technology).
- Wubbels, T., Creton, H. A., & Hooymayers, H. P. (1985, March). *Discipline problems of beginning teachers: Interactional teacher behaviour mapped out*. Paper presented at the annual meeting of American education research association.
- Wubbels, T., Creton, H., Levy, J., & Hooymayers, H. (1993). The model for interpersonal teacher behaviour. In T. Wubbels & J. E. Levy (Eds.), *Do you know what you look like? Interpersonal relationships in education* (1st ed., pp. 13–28). The Falmer Press.
- Yang, X. (2015). Rural junior secondary school students' perceptions of classroom learning environments and their attitude and achievement in mathematics in West China. *Learning Environments Research*, 18. <https://doi.org/10.1007/s10984-015-9184-3>

Adit Gupta is the Principal of MIER College of Education, Jammu. He has a Ph.D. in the field of 'Learning Environments' from Curtin University, Perth, Australia and master's degree in Psychology and Education. Dr. Gupta has over 26 years of teaching and professional experience at various levels. He is the recipient of the prestigious 'Endeavour Executive Award' of the Australian Government. Psychosocial Learning Environments, Educational Technology and Teacher-Student Interactions are his primary areas of research. He is also the Managing Editor of MIER Journal of Educational Studies, Trends and Practices. ORCID: <https://orcid.org/0000-0003-0018-608X>; email: adit@mier.in

Priya Sharma is a Research Scholar at MIER College of Education and is pursuing her M.Phil. in Education. She has a double postgraduate degree in Chemistry and Education. Priya is currently teaching in the Post Graduate department of Education at MIER College and overseeing an ICSSR funded research project. Her areas of interest are Science Education, Learning Environments and Teacher Education. email: priya.1905001@miercollege.in

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 13

Teaching Effectiveness in Spain: Towards an Evidence–Based Approach for Informing Policymakers



Carmen-María Fernández-García, Mercedes Inda-Caro,
and María-Paulina Viñuela-Hernández

Abstract This chapter uses a three-stage process of documentary analysis to illustrate how teaching effectiveness is assessed and studied in Spain. We begin by presenting the Spanish legal context, giving a historical overview of the most important education legislation. This is important, as there have been several reforms over recent years, and because of the decentralized model in Spain, which means that competencies and responsibilities are split between the Ministry of Education, Culture, and Sport and the regional administrations in the autonomous communities.

The second part focuses on educational innovations and effective teaching behaviors resulting from policy changes and the traditional and dominant paradigms in the Spanish educational landscape. We use the six teaching effectiveness domains from the ICALT project as a reference: safe learning climate, efficient classroom management, clarity of instruction, activating teaching, teaching-learning strategies, and differentiation. The third part describes empirical research undertaken in three autonomous communities in Spain to assess teaching quality. We look at organizational, human, and curricular factors which can help in interpreting teaching standards and their impact on student engagement.

Finally, the conclusions from the research are considered and discussed in terms of potential policy recommendations and practical decisions at both regional and national levels about teachers' initial training, continued training, and professional development.

Keywords Teacher effectiveness · Teacher behaviour · Student engagement · Questionnaires · Observation

C.-M. Fernández-García (✉) · M. Inda-Caro · M.-P. Viñuela-Hernández
Department of Educational Sciences, University of Oviedo, Oviedo, Spain
e-mail: fernandezcarmen@uniovi.es; indamaria@uniovi.es; paulina@uniovi.es

The social origins of the Spanish education system has meant the extension of the right to an education to all social groups and all levels. This is a consequence of Spanish society's awareness of the importance of education and the increasing expectations placed on it. With this in mind, the first part of this chapter considers this historical perspective in order to put the changes Spanish teachers have faced into context and make it understandable for international readers. The second part focuses on educational innovations and effective teaching behaviors resulting from policy changes and the traditional and dominant paradigms in the educational landscape in Spain. The third part describes empirical research which leads on to the final section, policy recommendations and practical implications at both national and regional levels.

1 Background of the Legal Framework of the Spanish Education System: Considering the Past to Understand the Present

The current Spanish Constitution was approved in 1978, and established a model of a decentralised state in which educational competencies were spread between all levels of government (Puelles, 1996). It was a symmetrical model in which the administrations in the autonomous communities had basically the same educational powers (Eurydice, 2021). Nowadays, the Spanish Ministry of Education, Culture and Sport (central government) establishes the fundamental rules of education (Blanch, 2011; Martínez-Usarralde, 2015). Central government ensures a common basic level in educational services, a coherent education system, and the equity of all citizens in educational terms (Aragon, 2013; Saenz, 2021). The autonomous administrations perform executive functions (Puelles, 1996); in other words, they apply these national regulations in their territories, as long as the application complies with the minimal teaching content established by central government, ensuring that there is a single educational system in Spain (in terms of its main features) (García, 2015). Finally, local authorities are responsible for the provision, repair, and maintenance of buildings and for ensuring school attendance where it is compulsory. The funding of the Spanish educational system also reflects this multi-level arrangement: autonomous administrations have assumed stewardship of educational spending, combining their own funds with money provided by the central government (Saenz, 2021).

There have been several reforms of the Spanish education system over recent decades. Although some of the reforms included significant changes related to parents' rights to choose the kind of school they wanted, and the participation of the educational community in education—*Ley Orgánica del Derecho a la Educación* (L.O.D.E.) [Right to Education Act] in 1985; *Ley Orgánica de Participación, Evaluación y Gobierno de los Centros Educativos* (L.O.P.E.G.C.E.) [Participation, Evaluation and Governance of Educational Institutions Act] in 1995—, only a few,

specific laws have changed the structure of the system or internal aspects of educational activity. These laws also changed the profile of students which had a clear impact on teacher behaviour and teaching methodologies.

The 1970 *Ley General de Educación* (L.G.E.) [General Education Act] made only primary education (*Enseñanza General Básica*) mandatory (ages 6–14). At 14 years old, students had to choose between vocational education and training (VET) [*Formación Profesional* (FP)] or academic upper-secondary education [*Bachillerato Unificado Polivalente* (BUP)] and then preparation for university [*Curso Orientación Universitaria* (COU)]. This last academic option was preferred by students who had good results or who had the firm intention to study at university. The social image, prestige, and expectations related to the two pathways were consequently very different (Carabaña, 1996).

The 1990 *Ley de Ordenación General del Sistema Educativo* (L.O.G.S.E.) [General Organization of the Education System Act] made it compulsory for students to stay in school until they were 16 years old. This meant that compulsory education consisted of primary education (from 6 to 12 years old) and compulsory secondary education [*Educación Secundaria Obligatoria* (ESO)] from 12 to 16 years old. Upper secondary education [*Bachillerato*] lasted two years, to 18 years old, and like access to vocational education and training, required students to have the certificate of compulsory secondary education.

Despite being short-lived, the 2002 *Ley Orgánica de Calidad de la Educación* (L.O.C.E.) [Quality of Education Act] included some measures changing the conception of academic achievement, for example by requiring students to repeat a year if they failed a certain number of subjects. Four years later, the *Ley Orgánica de Educación* (L.O.E.) [Education Act] emphasized dealing with individual needs and defined a more flexible education system, highlighting the need to facilitate the transition between educational stages.

The 2013 *Ley Orgánica para la Mejora de la Calidad Educativa* (L.O.M.C.E.) [Improvement in the Quality of Education Act] made slight changes to the structure of the final year of compulsory education. It established two options—academic and applied—in place of the previous arrangement, in which all students finished compulsory education following comprehensive programs. The act also included requirements to test before awarding certificates of compulsory secondary education and upper secondary education. Despite being part of the legislation, social opposition to these measures, which were felt to be segregational, made them difficult to apply.

In 2020, the Spanish government proposed a new reform of the education system with the *Ley Orgánica por la que se modifica la Ley Orgánica 2/2006 de Educación* (L.O.M.L.O.E.) [Modification of Education Act 2/2006], removing the dual option in the final year of compulsory education, removing final external exams, and adding a new branch of upper secondary education combining the sciences and humanities. Education in civics and ethics was given a larger role, focusing on human rights, sustainability and equity. Nevertheless, as before, cross-party agreement about education was again not possible.

Other reforms have also affected teachers' training (Viñao, 2013). Since 2010, teachers in secondary education and vocational education have had to have a relevant four-year university degree (*Grado*) and a master's in teacher training (Master's Degree in Teacher Training in Secondary and Upper Secondary Education and Vocational Training). This reform prioritized didactic and pedagogical factors which may contribute to improved teacher effectiveness. Nevertheless, there is yet to be a systematic assessment of the consequences of these changes, and there have been few studies about evaluating teaching effectiveness in Spain, especially outside higher education (Fernández-García et al., 2019; Herradas, 2021).

2 A Modern Conception of Teaching Effectiveness in Schools. Peculiarities of the Spanish Context

The model of teaching effectiveness behind the ICALT 3 project (International Comparative Analysis of Learning and Teaching) is based on six main domains allowing teaching tasks to be understood and executed (Van de Grift, 2007): safe learning climate; efficient classroom management; clarity of instruction; activating teaching; teaching-learning strategies; and differentiation. These domains outline a non-traditional concept of education. In this new concept, the student is the protagonist and this means that teachers have to employ complex strategies which match student learning styles and the paces at which they learn (Chocarro et al., 2007; Imbernon, 2012). This approach is a significant contrast to the traditional Spanish educational system, so it will be interesting to determine whether recent regulations fit in with this new concept of education. To that end, we look at each of the six teaching effectiveness domains, examining how they can be interpreted and viewed in the Spanish context.

2.1 Safe Learning Climate

A respectful safe learning climate is achieved when emotional and social intelligence go together. This promotes perseverance, management of impulsivity, use of a sense of humor, and the capacity to think independently (Costa & Kallick, 2008; Lucas & Claxton, 2014).

In this regard, when Montessori (n.d.) refers to the space or the classroom environment as a sign of respect for childhood in the *Casa dei Bambini* or when we analyze the school of Reggio Emilia we find a prepared environment that is safe, friendly, and full of stimuli, indicating that ethics must accompany aesthetics (Hoyuelos, 2004). The slow school movement (Holt, 2002; also see Quiroga, 2019) also considers this framework when it mentions the importance of studying in a

relaxed way, thoroughly covering each of the topics, and establishing relationships between knowledge and learning to think.

In this regard, Spanish schools must also ‘educate time’ (Novo, 2010), giving students the opportunity to be part of an environment which respects their needs and promotes comprehensive, integrated learning. For example, many schools (particularly public schools) do not stop for lunch (which is commonly eaten at 3 pm in Spain), so that children can finish their school day before they eat. A safe learning climate avoids excessive extracurricular activities; in class, students participate in the definition of activities so they understand what they are doing and why; and once activities are finished, there is time to review results with students. All of these examples contribute to creating a climate in which good relationships promote learning and in which students can combine academic, social, and personal learning.

2.2 *Efficient Classroom Management*

Concepts such as “slow pedagogy” (Holt, 2002; and see Quiroga, 2019) and “serene pedagogy” (Ritscher, 2013) reinforce the need for students to practice and learn to use time. Nowadays, this approach to using time removes the tension between time and syllabuses, preferring well-designed activities which facilitate the teaching-learning process.

One example of a way to achieve efficient classroom management is provided by the current *Programas de Diversificación Curricular* [Program of Curricular Diversification], with alternative ways of organizing timing and subjects (such as two-hour blocks in timetables rather than the traditional one hour and combining more than one subject in a single period): “*In this case, the objectives and competences will be achieved with a specific methodology organizing the curriculum in knowledge areas, practical activities and even different subjects* (article 27, L.O.M.L.O.E.). Another example is the problem-solving based methodologies used in some schools. In these non-traditional contexts, students have clearer ideas about what they are doing and why.

2.3 *Clarity of Instruction*

One of the main tasks of a teacher is to remove obstacles from the student’s path so that they can lead their own development. In this regard, the process used to gain knowledge is much more important than the knowledge itself (Steiner, 1961). Contemporary Spanish education has usually suffered from content overload; clarity of instruction requires selection and prioritizing relevant tasks. This is the only way to activate psychological capacities which will emerge from conversations, debates, and reflection. According to several authors (Domènech, 2009; Domènech

& Honoré, 2010; Honoré, 2006; Pastore, 2017; Thouless, 2017) the educational activities that are selected should define the time and not vice versa.

As part of this clarity of instruction, some Spanish schools are working holistically, following project-based learning methodologies so that students are encouraged to be more involved in their learning. New state-funded secondary schools have been designed with this idea in mind, meaning that the physical spaces, the teachers, and even the school timetables have been selected according to this paradigm.

2.4 *Activating Teaching*

The transition from information to knowledge needs relational learning and students have to be able to link their learning with their life stories so that they can perceive reality with new eyes (Esteve, 1983, 2010; Ventura, 2013). With this educational approach, teachers foster student curiosity and the “pedagogy of surprise” (Dewey, 1993; L’Ecuyer, 2013), and gain space to emphasize cognitive development (Melgarejo, 2013; Vygotski, 1998). There are also hybrid methodologies in this domain, which combine new and traditional techniques and usually produce better results in terms of academic results and student motivation (González-Marcos et al., 2021; Prieto et al., 2021).

In this sense, local education authorities in some autonomous communities are making significant efforts to install “dynamic classrooms”. They want to encourage alternative ways of organising learning spaces and stimulate the use of active methodologies including using information and communication technologies (I.C.T.) through flexible learning spaces (Educastur, 2021). These proposals are part of the Future Classroom Network promoted by the *Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado* (I.N.T.E.F.) [National Institute of Educational Technologies and Teacher Training].

2.5 *Teaching-Learning Strategies*

Teachers are urged to use a wide variety of teaching strategies. Traditional quality indicators will need to be reviewed and new assessment procedures are expected. In this sense, rankings of final results in different countries cannot be the sole reference as they do not take into account processes (Zavalloni, 2010, 2011). In the Spanish educational system, traditional classes have focused on telling students how things must be done. Nowadays teachers are developing other strategies such as letting students explain the processes needed to complete tasks or promoting knowledge exchange between students (Muelas, 2014). Interactive instruction will allow students to exercise control of teaching and learning processes, allowing them to reflect

about their learning and promoting “situated learning” (Hernández & Ventura, 2008) in which students become expert learners (Carnell & Lodge, 2002).

2.6 Differentiation

As pedagogies for inclusion and cooperation have indicated, human diversity must not be thought of as a problem, but rather an opportunity to reinforce individuals’ exceptionality and specificity (Skliar, 2017). Barbiana’s classic proposal pushed towards this way of understanding learning, avoiding labelling students by their grades and avoiding a rigid concept of the curriculum (Alumnos Escuela Barbiana, 1996; also see Carbonell, 2016). In a similar sense, Freinet (1978) suggested a kind of teaching and learning which considered education as a human right that can deal with social differences and diversity. Therefore, students will need different amounts of time for learning because of the paces they learn at, their needs, and their socio-cultural and family backgrounds.

The main strategies in Spain for improving teaching practices in terms of differentiation are considering students’ real levels of learning, pursuing significant learning and, as the most recent education legislation and regulations emphasize, addressing students’ special needs. The Spanish context is also diverse, which is reflected in the types of students and families and their educational expectations. There is broad variation between autonomous communities in, for example, the numbers of immigrant students (Instituto Nacional de Estadística, 2020), the proportion of private schools (Pérez et al., 2019), and the levels of school dropout (Ministerio de Educación y Formación Profesional, 2021). These examples indicate the different kinds of measures and teaching practices schools will need in order to deal with that diverse range of needs and requirements. *Aulas de Inmersión Lingüística* [Linguistic Immersion Classrooms], *Secciones Bilingües* [Bilingual Sections], *Programas de Diversificación Curricular* [Curricular Diversification Program], and *Formación Profesional Básica* [Basic Vocational Education] which were established by L.O.G.S.E. and reinforced in L.O.E., L.O.M.C.E. and L.O.M.L.O.E. are excellent examples of this.

3 Teaching Effectiveness in Spain: Contextual, Human, and Curricular Factors that Promote Better Teaching Skills

We now shift focus to explaining some of the key factors and variables that can help us understand teaching quality in the Spanish educational system.

The ICALT assessment instruments are validated tools that can be used to interpret and understand educational processes in schools. Given the lack of systematic

evaluation studies in the Spanish context, ICALT provides useful data allowing conclusions to be drawn about priorities and urgent needs in Spain. Although the ICALT sample was drawn from only three autonomous communities (and so cannot be used to generalize, merely indicate the specific patterns from that study), the results indicate that Spanish students generally feel that their teachers have appropriate skills in terms of learning climate, efficient classroom management, and instructional clarity. The six teaching effectiveness domains noted previously also have a significant relationship with student engagement (Fernández-García et al., 2019), a broad concept related to student behaviour and emotions related to dealing with academic tasks (Skinner et al., 2009). Despite that, students think that their teachers do not use enough active methodologies or a sufficiently wide range of teaching-learning strategies (Fernández-García et al., 2019); it seems that more innovative methodologies and greater use of ICT are expected. The recent pandemic and the prolonged impossibility of in-person teaching/learning underscored the need to improve this. Nevertheless, even teachers who had reported concerns or a lack of motivation about introducing these technological resources (Martín-Lucas et al., 2021) were able to achieve significant methodological transitions in a short time.

Spanish research has also shown that teachers suffer from high levels of social stress and face the challenge of dealing with student diversity when providing their students with up-to-date significant learning as well as developing students' skills to maintain an attitude of life-long learning (Gargallo et al., 2020). They also have to deal with a lack of resources and the social pressure resulting from continual changes in education legislation and hence the need to adapt to new social, financial, technological, and political conditions (Martínez-Otero, 2003; Pinel-Martínez et al., 2019; Viñao, 2004). A lack of rewards, and a perception of little social support help to explain anxiety disorders such as depression and 'burnout syndrome' (Doménech & Gómez, 2010; Silvero, 2007) which do not help teaching effectiveness. Lower and upper secondary education teachers are particularly affected by this issue (Pinel-Martínez et al., 2019). Reducing this anxiety and helping teachers needs us to look more deeply into teaching contexts and all of the internal and external elements that affect them.

3.1 Contextual Factors and Teacher Teaching Skills

As noted above, each of the 17 Spanish autonomous communities have to apply the general regulations to their territories. Although one might expect this to lead to differences that would make geographical location an important variable for teaching effectiveness, the data do not indicate significant differences in terms of teaching effectiveness between the three Spanish autonomous communities considered by ICALT (Inda-Caro et al., 2021). This may reflect central government's role in providing a unified, coherent educational system and future research should broaden its sampling to include participants from more of the country.

In contrast, results have indicated interesting differences depending on the type of schools. In this regard, there needs to be more detailed study of school-level policies and better understanding of schools' cultural contexts. This will help provide better interpretation of differences, given that teachers are encouraged to implement curricula that respect social and cultural diversity and are connected with the local experience.

Focusing on educational levels, Spanish students perceived better skills in lower secondary education teachers than teachers in upper secondary education or vocational education and training (Fernández-García et al., 2019). A more detailed examination of the variation by educational level would need separate consideration of each of the teaching skill domains.

3.2 Human Factors: Gender and Teaching Experience

Perceptions of teaching skills in Spain are affected by the gender of the teacher. The Spanish students in the ICALT study reported female teachers as having better skills in most of the teaching effectiveness domains, with the largest differences in differentiation strategies. Students thought that female teachers more clearly considered students' initial levels, produced more significant learning in their students, were better at checking whether students understood, and had a more realistic picture of students' difficulties in learning (Fernández-García et al., 2019). It seems that female teachers' views of education and student needs are a better fit with the demands of teaching effectiveness.

The results also indicated differences according to gender and educational level. Lower secondary students rated female teachers more highly than their male colleagues in clarity of instruction, activating teaching, differentiation, and teaching-learning strategies. In upper secondary education and vocational education and training, female teachers were perceived as better in teaching-learning strategies and efficient classroom management. Students in vocational education and training reported that female teachers paid more attention to differentiation strategies (Fernández-García et al., 2019).

Teaching experience was also found to affect the influence of teaching skills on student engagement (Inda-Caro et al., 2019) and the interaction of teaching experience and gender also played an important role. The Spanish students in the ICALT sample reported that male teachers with more teaching experience were less effective in their skills related to learning climate and efficient classroom management, whereas more experienced female teachers were seen as better in teaching-learning strategies such as prompting to summarize, giving strategies to learn new knowledge, and planning new ways to deal with novel tasks.

3.3 *Curricular Factors and Teaching Skills*

The curriculum is not monolithic, subjects are key components and are fundamental for understanding the teaching procedure. Based on our published results, teaching skills do not exhibit the same influence on student engagement in different subjects (Inda-Caro et al., 2021). Student gender also needs to be considered since it moderates this relationship.

For girls, the relationship between teaching behaviour and student engagement was stronger in the arts and physical education, particularly in terms of behavioural engagement. For emotional engagement, there were stronger relationships in exact and applied sciences. These findings from the more technical and scientific areas are particularly interesting because they underscore the teacher's role in increasing girls' enjoyment, self-assurance, interest, and involvement in subjects such as mathematics, physics, and computing. This remains vitally important as current studies have shown that women are not equally represented in the STEM [science, technology, engineering and mathematics] sector and perceive less support (Inda-Caro et al., 2017). Having identified this challenge, several proposals have been put forward in Spain to increase the presence of women in these areas (BBVA Research, 2017).

Looking at male students, the results showed that teaching skills had a stronger relationship with behavioural engagement in language, and vocational education and training subjects. For emotional engagement, teaching skills demonstrated stronger influence in social sciences and languages.

Vocational education and training (VET) subjects need particular attention. Girls' behavioural engagement in these subjects showed signs of greater improvement than in language, exact/applied sciences, or social sciences. However, in boys this effect occurred in emotional engagement (Inda-Caro et al., 2021). These different patterns show that teachers' tutoring roles should be very specific and appropriate for vocational education and training programs. In this regard, the classroom climate seems crucial in certain specialities that have traditionally been masculine teaching and learning spaces and girls should be encouraged to take on more active roles.

4 Practical Implications: Teacher and Student Roles, Two Key Factors for Improving the Teaching-Learning Process

Teachers and students are the key figures at either end of the teaching-learning process and both play a fundamental part in achieving a suitable emotional and motivational climate in the classroom. Students', teachers' and observers' perceptions of the emotional and motivational climate in the classroom, along with other teaching skills may help guide educational decision- and policymaking.

Spain needs to continue changing traditional teaching strategies. The concept of an educational “system” reinforces this idea because changes in any of the elements (teachers) necessarily means transformations in all the others (students’ relationships, internal organization of the classroom and so on). Spanish society and educational demands have changed enormously and educational processes must embrace these changes. Over the last twenty years, the continual reforms to the education system have obliged Spanish schools to try and establish alternative ways of understanding the teaching-learning process. In that changing Spanish educational context, teachers need more support from the authorities and those in charge of their professional development so that they feel more secure, especially in domains in which they feel there is room for improvement (e.g. clarity of instruction, activating teaching, teaching-learning strategies, and differentiation). There have already been improvements to the initial training that teachers receive, and perhaps now continued training and development should be the focus. This may improve the possibility of connecting fundamental and applied research and therefore exploring the full potential of not only initial teacher training, but also the support teachers need once they are working. This training should be focused on providing more teaching resources and pedagogical techniques, as well as on improving the psychological skills teachers need in order to cope with social and professional stress (Esteve, 1994; Hernández et al., 2020; Peñaherrera et al., 2014; Vicente & Gabari-Gambarte, 2019).

Giving teachers a clear picture of what they are expected to do and the precise behaviours which may help to improve student engagement would also make them feel more secure and relaxed, and help avoid unnecessary distress. In this regard, resource centres for training working teachers in the different autonomous communities may be fundamental (e.g. *Centros de Profesores y Recursos* (CPR) in Asturias, Extremadura, Murcia; or the *Centros de Profesorado* (CEP) in Andalucía, Cantabria, the Canary Islands, and the Balearics).

The role of teachers as professionals within society also needs to be strengthened, highlighting their qualifications and attempting to reassert the positive reputation that teachers and teaching had at the beginning of the twentieth century. Cross-party agreement about education would also give teachers a more stable environment and greater consensus about access to the teaching profession. These measures will help clarify teachers’ social image and provide a clearer definition of their professional competencies, distinguishing those professional skills from other “social” competencies which have contributed to teachers’ high workload (Esteve, 1994; Llorens et al., 2003). The pandemic may have helped to emphasize how important teachers’ roles are, as they worked to keep their students involved with learning tasks and to avoid leaving any children behind (López, 2021).

The ICALT project gives us interesting information that helps identify the most important domains in order to direct changes towards improving student engagement. This is essential, because research has shown that student engagement determines motivation and achievement and reduces the risk of dropout and school failure (Finn, 1989; Fredricks et al., 2011; Opdenakker & Minnaert, 2011; Skinner & Belmont, 1993).

In summary, there needs to be a deep understanding of the foundations and the theoretical principles of education and teaching activities which can guide policy-makers, researchers, and teachers in interpreting and understanding the practical results of research. The ICALT project responded to the lack of systematic procedures for teacher assessment in Spain, giving information based on the opinions of teachers, students, and external observers, resulting in a valid model for assessing the best direction for future changes. ICALT allowed a single instrument to be used to analyse teaching practice along with the possibility of interpreting the results according to the particular conditions in the different parts of Spain. This dual approach is the only way to guide changes securely, based on the evidence. Contextual, human, and curricular factors provide significant pointers towards the actions needed to improve teaching practice and therefore student engagement in Spain.

References

- Alumnos Escuela de Barbiana. (1996). *Carta a una maestra*. PPC.
- Aragón, M. (2013). Las competencias del Estado y las comunidades autónomas sobre educación. *Revista Española de Derecho Constitucional*, 98, 191–199.
- BBVA Research. (2017). *El futuro del empleo*. BBVA, CES Castilla y León.
- Blanch, J. (2011). Descentralización y autonomía en el sistema educativo en España. El caso de Catalunya. *Italian Journal of Sociology of Education*, 2, 11–30.
- Carabaña, J. (1996). ¿Se devaluaron los títulos? *Revista Española de Investigaciones Sociológicas*, 75, 173–213.
- Carbonell, J. (2016). *Pedagogías del siglo XXI. Alternativas para la innovación educativa*. Octaedro.
- Carnell, J., & Lodge, C. (2002). *Supporting effective learning*. Pau Chapman.
- Chocarro, E., González-Torres, M. C., & Sobrino, A. (2007). Nuevas orientaciones en la formación del profesorado para una enseñanza centrada en la promoción del aprendizaje autorregulado de los alumnos. *Estudios sobre educación*, 12, 81–98.
- Costa, A. L., & Kallick, B. (Eds.). (2008). *Learning and leading with habits of mind: 16 essential characteristics for success*. ASCD.
- Dewey, J. (1993). *Cómo pensamos. Nueva exposición de la relación entre el pensamiento reflexivo y proceso educativo* (Original 1933). Paidós.
- Domènech, J. (2009). *Elogio de la educación lenta*. Graò.
- Domènech, F., & Gómez, A. (2010). Barriers perceived by teachers at work, coping strategies, self-efficacy and burnout. *The Spanish Journal of Psychology*, 13(2), 637–654. <https://doi.org/10.1017/S1138741600002316>
- Domènech, J., & Honoré, C. (2010). Romper el tabú de la lentitud. *Aula de innovación educativa*, 193, 20–22.
- Educastur. (2021). *Aulas dinámicas*. <https://www.educastur.es/aulas-dinamicas>
- Esteve, J. M. (1983). El concepto de educación y su red nomológica. In J. L. Castillejo et al. (Eds.), *Teoría de la Educación I. El problema de la educación* (pp. 9–25). Ediciones Límites.
- Esteve, J. M. (1994). *El malestar docente*. Paidós Ibérica.
- Esteve, J. M. (2010). *Educación, un compromiso con la memoria: un libro para educar en libertad*. Octaedro.
- Eurydice. (2021). *Organization and governance of the Spanish Educational System*. https://eacea.ec.europa.eu/national-policies/eurydice/content/organisation-and-governance-79_en

- Fernández-García C. M., Maulana R., Inda-Caro M., Helms-Lorenz, M., & García-Pérez, O. (2019). Student perceptions of secondary education teaching effectiveness: General profile, the role of personal factors, and educational level. *Frontiers in Psychology, 10*, 533–543. <https://doi.org/10.3389/fpsyg.2019.00533>
- Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research, 59*, 117–142.
- Fredricks, J., McColskey, W., Meli, J., Mordica, J., Montrosse, B., & Mooney, K. (2011). Measuring student engagement in upper elementary through high school: A description of 21 instruments. *Issues and Answers Report, 98*, 1–80.
- Freinet, C. (1978). *Las técnicas Freinet de la escuela moderna*. Siglo veintiuno.
- García, J. (2015). El proceso de descentralización educativa en España. *Edetania, 48*, 203–216.
- Gargallo, B., García-García, F. J., Moreno, S., López, I., & Jiménez, M. A. (2020). La competencia aprender a aprender: valoración de un modelo teórico. *Revista Española de Pedagogía, 276*, 187–211. <https://doi.org/10.22550/REP78-2-2020-05>
- González-Marcos, A., Navaridas-Nalda, F., Jiménez-Trens, M. A., Alba-Elías, F., & Ordieres-Meré, J. (2021). Efectos académicos de una enseñanza mixta versus metodología única centrada en el profesor y enfoques de aprendizaje. *Revista de Educación, 392*, 123–143. <https://doi.org/10.4438/1988-592X-RE-2021-392-481>
- Hernández, F., & Ventura, M. (2008). *La organización del currículum por proyectos de trabajo*. Octaedro.
- Hernández, M. A., Panalva, A., & Guerrero, C. (2020). Profesorado y convivencia escolar: necesidades formativas. *Magister, 32*(1), 23–32. <https://doi.org/10.17811/msg.32.1.2020.23-32>
- Herradas, S. (2021). Evaluación de la práctica de los docentes de educación obligatoria, estudio comparado en Europa. *Revista Española de Educación Comparada, 39*, 238–253. <https://doi.org/10.5944/reec.39.2021.27388>
- Holt, M. (2002). It's time to start the slow school movement. *Phi Delta Kappan, 84*(4), 264–271.
- Honoré, C. (2006). *Elogio a la lentitud*. RBA.
- Hoyuelos, A. (2004). *La ética en el pensamiento y la obra de Loris Malaguizzi*. Octaedro.
- Imberón, F. (2012). Nuevos retos del profesorado en el siglo XXI: la necesidad de una nueva formación permanente en tiempos de crisis. *Crítica, 982*, 24–27.
- Inda-Caro, M., Rodríguez-Menéndez, M. C. & Torío-López, S. (2017). Evaluación del modelo cognitivo social de desarrollo de la carrera para la predicción de las metas en las materias tecnológicas de estudiantes de bachillerato. *Estudios sobre Educación, 32*, 49–71. <https://doi.org/10.15581/004.32.49-71>
- Inda-Caro, M., Maulana, R., Fernández-García, C. M., Peña-Calvo, J. V., Rodríguez-Menéndez, M. C., & Helms-Lorenz, M. (2019). Validating a model of effective teaching behaviour and student engagement: Perspectives from the Spanish students. *Learning Environment Research, 22*, 229–251. <https://doi.org/10.1007/s10984-018-9275-z>
- Inda-Caro, M., Fernández-García, C., Maulana, R., & Viñuela-Hernández, P. (2021). The effect of contextual, personal and curricular factors on students' engagement. *Revista de Educación, 391*, 181–198. <https://doi.org/10.4438/1988-592X-RE-2021-391-474>
- Instituto Nacional de Estadística. (2020). *Saldo migratorio interautonómico*. <https://www.ine.es/jaxiT3/Datos.htm?t=24378#!tabs-tabla>
- L'Ecuycer, C. (2013). *Educar en el asombro. ¿Cómo educar en un mundo frenético e hiperexigente?* Plataforma Actual.
- Ley Orgánica 3/2020, de 29 de diciembre, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación (L.O.M.L.O.E.). *Boletín Oficial del Estado, 340*, 30 de diciembre de 2020, 122868–122952.
- Llorens, S., Salanova, M., & García-Renedo, M. (2003). ¿Por qué se están quemando los profesores? *Prevención, trabajo y salud, 28*, 16–24.
- López, F. (2021). La profesión docente en la perspectiva del siglo XXI. Modelos de acceso a la profesión, desarrollo profesional e interacciones. *Revista de Educación, 393*, 69–96. <https://doi.org/10.4438/1988-592X-RE-2021-393-486>
- Lucas, B., & Claxton, G. (2014). *Nuevas inteligencias, nuevos aprendizajes*. Narcea.

- Martínez- Otero, V. (2003). Estrés y ansiedad en los docentes. *Pulso: Revista de Educación*, 26, 9–22.
- Martínez-Usarralde, M. J. (2015). Spain. In W. Horner, H. Dobert, L. R. Reuter, & B. Von Kopp (Eds.), *The Education systems of Europe* (pp. 759–778). Springer.
- Martín-Lucas, J., Torrijos-Fincias, P., Serrate-González, S., & García del Dujo, A. (2021). Teaching use intention and self-perception of bLearning in higher education. *Revista de Educación*, 391, 199–224. <https://doi.org/10.4438/1988-592X-RE-2021-391-475>
- Melgarejo, X. (2013). *Gracias, Finlandia*. Laertes.
- Ministerio de Educación y Formación Profesional. (2021). *Datos y cifras. Curso Escolar 2021/2022*. <https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/indicadores/datos-cifras.html>
- Montessori, M. (n.d.). *Manual práctico del método*. Ediciones Araluce.
- Muelas, A. (2014). Competencias del profesorado de educación secundaria obligatoria y bachillerato. Construyendo desde las estrategias de aprendizaje. *International Journal of Developmental and Educational Psychology INFAD Revista de Psicología*, 1(7), 185–192. <https://doi.org/10.17060/ijodaep.2014.n1.v4.625>
- Novo, M. (2010). *Despacio, despacio: 20 razones para ir más lento por la vida*. Obelisco.
- Opdenakker, M. V., & Minnaert, A. (2011). Relationships between learning environment characteristics and academic engagement. *Psychological Reports*, 109(1), 259–284. <https://doi.org/10.2466/09.10.11.PR0.109.4.259-284>
- Pastore, F. (2017). Why so slow? The school-to-work transition in Italy. *Discussion Paper Series*, 10767, 1–39.
- Peñaherrera, M., Cachón, J., & Ortiz, A. (2014). La autoestima profesional docente y su implicación en el aula. *Magister*, 26(1), 52–58. [https://doi.org/10.1016/S0212-6796\(14\)70018-0](https://doi.org/10.1016/S0212-6796(14)70018-0)
- Pérez, F., Serrano, L., & Uriel, E. (Dir.). (2019). *Diferencias educativas regionales 2000–2016: Condicionantes y resultados*. Fundación BBVA.
- Pinel-Martínez, C., Pérez-Fuentes, M. C., & Carrión-Martínez, J. J. (2019). Investigación sobre el *Burnout* en docentes españoles: una revisión sobre factores asociados e instrumentos de evaluación. *Bordón*, 71(1), 115–131. <https://doi.org/10.13042/Bordon.2019.62122>
- Prieto, A., Barbarroja, J., Álvarez, S., y Corell, A. (2021). Eficacia del modelo de aula invertida (flipped classroom) en la enseñanza universitaria: Una síntesis de las mejores evidencias. *Revista de Educación*, 391, 149–177. <https://doi.org/10.4438/1988-592X-RE-2021-391-476>
- Puelles, M. (1996). Educación y autonomía en el modelo español de descentralización. *Revista de Educación*, 309, 163–193.
- Quiroga, P. (2019). La educación lenta en perspectiva histórica: conceptualización, desarrollo y concreción en las iniciativas de Madre de día Waldorf en España. *Educació i Història: Revista d'Història de l'Educació*, 34, 121–147. <https://doi.org/10.2436/20.3009.01.232>
- Ritscher, P. (2013). *Escola slow. Pedagogia del quotidià*. Rosa Sensat.
- Saenz, E. (2021). La financiación de la educación en el Estado autonómico y su coherencia con la distribución de competencias educativas. *Revista de Derecho Político*, 111, 45–76. <https://doi.org/10.5944/rdp.111.2021.31045>
- Silvero, M. (2007). Estrés y desmotivación docente: el síndrome del “profesor quemado” en educación secundaria. *Estudios sobre educación*, 12, 115–138.
- Skinner, E. A., & Belmont, M. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, 85(4), 571–581. <https://doi.org/10.1037/0022-0663.85.4.571>
- Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children’s behavioural and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement*, 69, 493–525. <https://doi.org/10.1177/0013164408323233>
- Skliar, C. (2017). *Pedagogías de las diferencias*. Novedades Educativas.
- Spanish Constitution [Constitución Española]. *Boletín Oficial del Estado*, 311, (29/12/1978) (legislado).

- Steiner, R. (1961). *Una introducción a la Educación Waldorf*. Rudolf Steiner.
- Thouless, M. D. (2017). Slow and steady. The effects of teaching a one-semester introductory mechanics class over a year. *The International Journal of Engineering Education*, 33(6), 1842–1855.
- Van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research*, 49(2), 127–152. <https://doi.org/10.1080/00131880701369651>
- Ventura, M. (2013). La perspectiva educativa de los proyectos de trabajo como motor de cambio en la organización del curriculum y en la gestión del tiempo y el espacio. *Investigación en la escuela*, 79, 19–30.
- Vicente, M. I., & Gabari-Gambarte, M. I. (2019). Burnout y Factores de Resiliencia en Docentes de Educación Secundaria. *International Journal of Sociology of Education*, 8(2), 127–152. <https://doi.org/10.17583/rise.2019.3987>
- Viñao, A. (2004). *Escuela para todos: educación y modernidad en la España del siglo XX*. Marcial Pons.
- Viñao, A. (2013). Modelos de formación inicial del profesorado de educación secundaria en España (siglos XIX–XX). *Revista Española de Educación Comparada*, 22, 19–37. <https://doi.org/10.5944/reec.22.2013>
- Vygotski, L. (1998). *El desarrollo de los procesos psicológicos superiores*. Editorial Planeta.
- Zavalloni, G. (2010). Por una pedagogía de caracol. *Aula de innovación educativa*, 193–194, 23–26.
- Zavalloni, G. (2011). *La pedagogía del caracol: por una escuela lenta y no violenta*. Editorial Graó.

Carmen-María Fernández-García, PhD, Associate Professor at the Department of Educational Sciences at the University of Oviedo (Spain). She has received research grants from the Spanish Ministry of Education. She is member of the Spanish Society of Comparative Education, the Spanish Society of Pedagogy and the ASOCED Research Group. Her major research interests involve teaching and teacher education, learning and instruction, gender and comparative education. She has published several academic papers on these topics. Currently she is joining an international project investigating teaching behavior and student outcomes across countries, the ICALT3 Project coordinated by the University of Groningen.

Mercedes Inda-Caro, PhD, Associate Professor at the University of Oviedo (Spain). She previously worked as a training support counselor in a public school as part of her FICYT scholarship training (1997) and as Child Educator for the Principality of Asturias within the Ministry of Social Services in two periods (1996/2000). Her PhD dealt with the concept of personality disorders. Currently, she is working on three lines of research: family and gender, teacher and teaching-learning education, and gender and technology studies, as a member of the ASOCED Research Group. She has several publications in scientific journals.

María-Paulina Viñuela-Hernández, PhD, Associate Professor at the Department of Educational Sciences at the University of Oviedo (Spain). She received her PhD, focusing on professional training and employment programs. She is member of the Iberoamerican Society of Social Pedagogy (SIPS). Her main research interests include teaching and instruction in the fields of occupational training, intercultural education, teachers' training gender and education. She has published several academic papers on these topics. Currently she is joining an international project investigating teaching behavior and student outcomes across countries, the ICALT3 Project coordinated by the University of Groningen.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 14

An Explanation of the ICALT Instrument's Measurement of Teaching Quality in Relation to Teacher Education and Policy in South Korea



Seyeoung Chun, Okhwa Lee, and Deuk-Joon Kim

Abstract The rapid development of South Korea's educational system has attracted international interest. The country is well-known for its high student achievement, as indicated by the OECD PISA research, yet the causes for the high achievement remain unclear. Many argue that high teacher quality is an explanatory variable, even though accurate and rigorous measurement of teaching quality at both the practical and theoretical levels has yet to be established. The ICALT (International Comparative Analysis of Learning and Teaching) developed by van de Grift and colleagues in the Netherlands was recently utilized to assess the teaching quality of Korean teachers, and the results demonstrated a high level of teaching quality when compared to other countries. In this chapter, we discuss the relationship between the ICALT's reported high level of teaching quality and teacher education and policy in South Korea. Several components of teacher education and policy are identified as factors that lead to the quality of the teaching force. They are the well-developed teacher training system, higher level of teachers' socioeconomic status, in- & external-school supervision for enhancing teacher competency, and efficient personal administration for teachers including homeroom teacher, rotation and promotion.

Keywords ICALT · Teaching quality · Teacher education · South Korea

S. Chun (✉)

Department of Education, Chungnam National University, Daejeon, South Korea
e-mail: sychun56@gmail.com

O. Lee

Department of Education, Chungbuk National University, Cheongju, South Korea
e-mail: ohleekorea@gmail.com

D-J. Kim

Chungnam National University, Daejeon, South Korea
e-mail: deukjoon@gmail.com

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_14

1 Introduction

Korea's rapid economic and social development during the last decades has been attributed to its educational success and development. Changes in the education system have been remarkable in both quantity and quality in the last 70 years. There are so many indicators of educational development that they are difficult to enumerate: for instance, almost 90% of the whole school-age population graduated from high school and entered the tertiary education system in the recent decades, and the illiteracy rate is drastically reduced down to less than 10% from more than 70% since 1945. Universal attainment of primary education was achieved in the 1960s and secondary education in the 1970s. In this chapter, we will explore the findings from Korean administrations of the ICALT (International Comparative Analysis of Learning and Teaching) measure, and analyze connections with Korean teacher education and policy.

One of the most compelling proofs of South Korea's educational power is the outstanding results in the various international assessment of student achievement in recent years. In the last PISA (Program for International Student Assessment) study conducted by the OECD (Organization for Economic Cooperation and Development) in 2018, Korean students were placed in the top tier category. According to the snapshot of South Korea from PISA 2018 country-specific overviews about "What 15-year-old students in Korea know and can do," Korean students scored higher than the OECD average in reading, mathematics, and science. Compared to the OECD average, a larger proportion of students in Korea performed at the highest levels of proficiency (Level 5 or 6) in at least one subject; at the same time, a larger proportion of students achieved a minimum level of proficiency (Level 2 or higher) in at least one subject.

However, little is known about how Korean success and development have been achieved. Quality of teaching is often selected as one of the most convincing factors. Few disagree that the quality of a teacher is the most important aspect of a student's academic success, as it is commonly stated that "the quality of education cannot exceed the quality of teachers." Much past research on student accomplishment has concluded that school disparities are ultimately due to teacher variations and that individual teachers, irrespective of schools, have a significant impact on pupils (Marzano et al., 2001). van de Grift et al. (2017) reviewed a substantial body of research regarding the relationship between teacher quality and student learning and summarized that the results of these research efforts made clear that about 15–25% of the differences in students' achievement might be explained by the work of teachers.

In this sense, many aspects related to the quality and quantity of the teaching force in South Korea can support the plausible reasons for the outstanding performance of students. The teachers in South Korea are selected from the best-talented people and are very well paid. All schools are evenly provided with those good teachers regardless of regional disparities due to the constitutional mandate that everybody has the right to equal education based on ability.

On the other hand, in order to ensure the good quality and quantity of the teaching force for the aim of quality education, government policy efforts have been significantly intensified. In that view, establishing the professionalism of the teaching job has been prioritized: that is, the teacher is entitled to be the expert, the professional who distinguishes themselves from ordinary and general employees. Although there can be many arguments about what it means in the reality of a teaching job or how it can be differentiated from other jobs, several researchers have classified teaching as a professional occupation (Flexner, 1910; Lieberman, 1956). The notable document that specifies the professionalism of teachers would be the ILO/UNESCO Recommendation concerning the Status of Teachers (1966). Article 6 of the Recommendation states, *‘Teaching should be regarded as a profession: it is a form of public service which requires of teachers expert knowledge and specialized skills, acquired and maintained through rigorous and continuing study; it calls also for a sense of personal and corporate responsibility for the education and welfare of the pupils in their charge’*. Article 31 (4) of the Constitution of the Republic of Korea and Article 14 of the Framework Act on Education also stipulate together that “the professionalism of teachers in school education is respected...”.

However, professionalism about the characteristics of the teaching job is very difficult to conceptualize at the academic level as well as the practical level. It is very different from a subjective teacher's point of view. According to a study on the reconceptualization of teacher expertise (Kim, 2006), a teacher's expertise or professionalism is defined as an individual teacher's ability to build skills through experience and training based on their beliefs and knowledge, and to perform the teaching profession appropriately in the school setting. Nonetheless, there are numerous classifications for the concept of teacher knowledge, and there are frequently disagreements and controversies when it comes to real-world issues. Despite these different considerations, there is a tendency to confine teachers' competence to classroom instruction and teaching. Among the many things a teacher performs including classroom teaching, student mentoring and counseling, and other various administrative affairs, classroom teaching is supposed to be at the heart of what a teacher does. Hence even the quality of a school itself may be measured by how classroom teaching is handled, which means the classroom teaching quality is at the heart of the teaching profession.

OECD-TALIS can be regarded as a sister study project to the PISA on students' achievement, started in 2008. According to the OECD/TALIS homepage, the Teaching and Learning International Survey (TALIS) is the first international survey that provides a voice to teachers and school principals, who complete questionnaires about issues such as the professional development they have received; their teaching beliefs and practices; the assessment of their work and the feedback and recognition they receive; and various other school leadership, management and workplace issues ([http:// https://www.oecd.org/education/talis/talisfaq/](http://https://www.oecd.org/education/talis/talisfaq/)). As indicated “it is not an assessment, but a self-reported survey,” The TALIS study focuses on the teaching quality as a kind of skill that can be assessed or measured, but is limited to reporting on the teachers' working conditions by their own voices.

According to the TALIS study, Korean teachers demonstrated lower levels of self-efficacy and job satisfaction, as indicated in Table 14.1. In the same TALIS report, it's also interesting to find Korean teachers' autonomy at a higher level, whereas Finnish teachers' autonomy is at a lower level. Finnish education and Korean education are often compared as they both have high students' performance yet the educational culture is known very different but the social status of teachers is similar in terms of social respect and economic rewards.

This raises the possibility that teachers' competence for teaching effectiveness may not be explained by teachers' self-efficacy, satisfaction, or autonomy in explaining where Korean students' excellent performance comes from. According to the TALIS study, teacher-related factors are not directly associated with teaching quality; rather, they are indirect variables that help teachers teach effectively. The search for a direct metric of teaching quality that can explain student success is thus worthwhile. In juxtaposition to their pupils' strong achievement, this negative or lower evaluation report from Korean teachers is a very interesting phenomenon. This phenomenon was stated as the 'Korean Paradox' by Kim et al. (2009a: 23–24): "There have been controversies over the role of teachers regarding the remarkable results of Korean students' achievements. Some critics argue that the academic success of many Korean students is due to private tutoring, rather than their classroom teachers. ...However, the government likes to claim that the Korean PISA achievements are a result of the outstanding educational system and teachers. In some sense, this might be true. ... It might be assumed that the high qualifications of Korean teachers are related to students' achievement in some ways, but solid empirical evidence is lacking to support this claim definitely."

This paradox arises from the lack of a firm foundation of knowledge upon which to evaluate educational quality. A recent research initiative called ICALT (International Comparative Analysis of Learning and Teaching) may provide a way out of this conundrum. The ICALT instrument has been demonstrated to be a scientific and accurate tool for measuring and comparing the quality of teaching in various countries and cultures. It was created in the Netherlands by Wim van de Grift and others. In this chapter, the findings of the ICALT instrument's assessment of

Table 14.1 Trend of change in teaching-learning efficacy (%)

Sorted	Makes good questions for students		Uses a variety of assessment strategies		Explains in different ways when students do not understand	
	TALIS 2013	TALIS 2018	TALIS 2013	TALIS 2018	TALIS 2013	TALIS 2018
Korea	77.4	86.6	66.6	78.0	81.4	89.7
Average of participating countries	87.4	86.7	81.9	81.0	92.0	91.6

Source: reconstructed data from OECD (2019). TALIS 2018 Results: Teachers and School Leaders as Lifelong Learners

Korean teachers' teaching quality will be presented and analyzed in connection to Korean teacher education and policy.

2 Teaching Quality of Korean Teachers

The reason for the disparity and scarcity of information on teaching quality in Korea is that there is no objective and accurate methodology for measuring teaching quality, i.e., we haven't had a good tool to illustrate how well teachers behave themselves in the classroom. Such information and statistics did not exist. However, various studies and approaches have lately been established to scientifically observe and quantify teaching quality and competencies.

Prior research on teacher behavior to improve teaching skills provided general rules and principles, helped to describe the phenomenon and helped to reveal the effectiveness of specific teaching behaviors, but a scientific approach to teaching behavior in the overall classroom context was still uncommon (Chun et al., 2017). In this regard, the research conducted by the van de Grift team at the University of Groningen in the Netherlands has consistently produced a number of positive results in this area by observing teachers' instructional behavior in the classroom which revealed the level of instructional skills, and providing feedback and coaching for improvement (van de Grift, 2007). The Dutch research team expanded it to the worldwide level and titled it ICALT, which stands for International Comparative Analysis of Learning and Teaching, based on various studies conducted in Europe with persuasive results.

This global application of ICALT research began in 2014 with the ICALT III project, in which 18 countries were involved: the Netherlands, Korea, Indonesia, the United Kingdom, China, Hong Kong, Spain, South Africa, Turkey, Malta, the United States of America, Norway, Australia, Nicaragua, Mongolia, Pakistan, Portugal, and Brazil. The study's main topic was whether the quality of teaching can be compared across countries in terms of reliability and validity. Several studies have been published in journals (Maulana et al., 2020a, b; Andre et al., 2020; van de Grift et al., 2017, 2019), demonstrating the reliability and validity of the ICALT observation tool. Those comparative ICALT studies were conducted for secondary school teachers in a few countries, and a comparison for all nations is not finished yet. Using this research instrument, however, it was demonstrated that the ICALT tool may be utilized for worldwide comparative research and that differences in teaching quality can be measured.

The ICALT tool was used for the first comparative study on the teaching expertise in Korea and Netherlands in 2014: 289 Dutch secondary school teachers and 375 Korean secondary school teachers participated. It was found that the six ICALT scales for measuring teaching skills, assessed in South Korea and the Netherlands, were sufficiently reliable and offered sufficient predictive value for student engagement. Multigroup confirmatory factor analysis showed that the factor loadings and intercepts of the six ICALT scales were the same, within acceptable boundaries, in

both countries. This means the average scores of teachers in both countries assessed by the tool can be compared in a reliable and valid way. According to the research, it was found that Dutch secondary teachers fared marginally better in the 1–4 categories of teaching skills, while Korean secondary teachers did better in more advanced teaching domains. In other words, Korean secondary school teachers outperformed Dutch secondary school teachers in the 5–6 domain, the most advanced levels. Provided that those advanced teaching skills have great potential to influence the learning gains of both struggling and excellent learners, it might also contribute, amongst other factors, to the higher level of student engagement evident from the first ICALT research findings in the South Korean sample. According to these findings, the reason why Korean students outperform Dutch students in OECD-PISA accomplishments could be attributed to the high level of teaching expertise in Korea.

Every year, the ICALT-K Korea Research Center (Chief: Seyeoung Chun, Professor of Chungnam National University) trains observation experts and conducts ICALT data collection through class observations of Korean elementary and secondary school teachers. ICALT-K Korea Research Center collected 1976 classroom teaching samples from 2014 to 2021; 598 elementary instructors, 936 middle, and 442 high school teachers; 539 male teachers, and 1420 female teachers. Since the experiment began, 72 trained observers have participated in the observation. They have been attending annual ICALT observation training given by the research center, and Cohen's kappa has shown that they have reached a satisfactory level of agreement of over .70. The statistical criteria for worldwide comparability were also found to have passed the reliability and validity test. The construct validity estimates for all 32 items ranged from .550 to .896, which is higher than the lower threshold of .5. The construct dependability of all six domains was over .90, and the variance extract index was over .60.

Figure 14.1 shows the descriptive level of teaching skills. Although there are slight differences by school level between elementary and secondary, the data leads us to conclude that Korean teachers display very high levels of teaching expertise.

In 2020, an international comparative study of secondary school teachers' teaching skills in six nations (the Netherlands, Korea, South Africa, Indonesia, Hong Kong, and Pakistan) was published, with Korean teachers scoring top in all disciplines (Maulana et al., 2020c). Results of the study showed that South Korean teachers were rated higher in all domains ($p < 0.001$), except for learning climate. Higher ratings on most of the teaching behavior domains for South Korean teachers compared to Dutch, South African, and Indonesian teachers might be related to several effective teaching supporting factors including how teachers in the country are recruited, how they value learning, and how they are supported professionally. There must be various factors reasoning for the high performance of the Korean education system. However, even though that reasoning sounds logical, it must be empirically validated. In this sense, ICALT approach for assessing and comparing the teacher quality and skills is worth valuing its contribution to a better understanding of the quality of teaching as a good factor of Korean success in education.

Based on the ICALT framework, it is plausible to assume that South Korean instructors retain a greater level of teaching expertise, quality, and skills, which

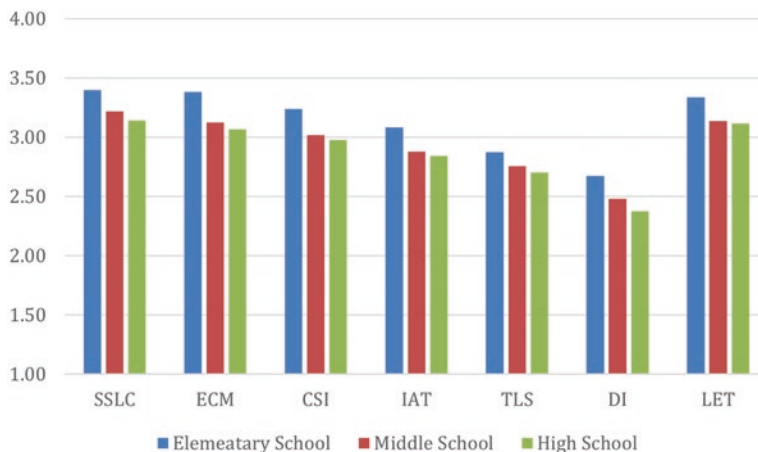


Fig. 14.1 Study on ICALT class expertise-average by school level. (Source: The graph was created by the authors using the data collected by the e-ICALT platformI (<http://icalt.kr>) which is the data collection site to which the trained observers upload the classroom observation data by ICALT tool within the framework of ICALT-K)

explains South Korean students' higher level of learning success in many international studies, such as the OECD/PISA. The following stage in the inquiry is to naturally point to the sources behind the excellent quality of teaching. As is generally known in Korea, there have previously been many suggestions regarding the sources. Teaching jobs have several attractive advantages for workers, such as high job stability, relatively high stable salary, social respect, and lifetime employment.

3 Teacher Education and Policy in South Korea

A number of factors contribute to Korean teachers' high level of teaching ability. However, there is a paucity of information on how Korean teachers can become world leaders in their teaching skill. Every year, the OECD research reveals that the teaching job in Korea is unquestionably one of the most attractive careers in the country, since teachers not only earn the highest compensation in the world, but also have a work guarantee until the age of 62, and have socially high prestige and respect as public servants. It wasn't always like this, though. Teachers' socioeconomic position was quite poor until the 1970s. In order to enhance teacher status and quality, various policy efforts and tools have been formulated and implemented over the years.

Teachers should have a national teaching license, which can be obtained from the four-year pre-service teacher training at colleges. Those who want to work as teachers after graduating from pre-service institutions compete for jobs at public schools by taking a demanding examination. After the completion of university education,

those graduates with the national teaching certificate should pass the national examination to be allocated to public schools, and before teachers start to work at schools, they need to take official training, the first in-service education. The remaining processes for a teacher's career are job assignments at schools as a public official, promotion to principal, and finally retirement at the age of 62 with the honorable award of the Order of Service Merit and retirement pension. It is vital to understand the teaching profession within the context of Korean educational policy in the past 70 years.

3.1 Pre-service Teacher Training

South Korea can attract high quality candidates for teacher education institutes. Both entering the university (school of education) and the recruiting test are very competitive. This is not the case in all countries: qualified personnel are in short supply in both developing and developed countries. In the United States, for example, a poll found that approximately 70% of teachers traditionally score below the national average on the SAT, a college entrance exam (Kim, 2006). However, Korea can recruit the very best high school graduates and this tendency has a long tradition from the start of the Korean education system in the 1960s and 1970s. In Korea, initially, there were not enough qualified teachers to meet the demands of education when the population grew rapidly so the demand for expanding the educated population was high.

Teacher Training System When the Korean education system first began, teacher training institutions were in short supply. Instructors' socioeconomic remuneration, including personal treatment and working circumstances, were exceedingly low at the start of Korea's public education system, which made it difficult to recruit qualified teachers. Teaching jobs may not achieve the degree of the economic standard that other jobs might offer during periods of high economic expansion. As a result, the government devised a scheme to entice young people by exempting them from military service and by paying their university tuition. Those initiatives, however, were insufficient to attract young and skilled workers. However, because of the *Sungmun* culture, high regard for the educated and academics, even when the labor market was constrained during a period of economic hardship, teaching jobs remained attractive to young people. As a result, the job market for teachers has returned as a supplier and producer of the education sector as a kind of booming industry, which has resulted in education expansion and development. As a result of a large number of candidates, teacher training colleges were invited to generate the teaching force. At the same time, the number of graduates was insufficient to meet the constantly growing number of schools and pupils. It was a kind of virtuous circle of supply and demand in the education sector.

Changes in the Teacher Training System *Hansung Normal School* began teaching elementary school teachers in 1895 and was promoted to a two-year college in 1961 as the national college of education. Initially, there were ten colleges, which were eventually increased to sixteen. From 1981 to 1984, they were converted to four-year bachelor's universities, and in 1993, they were renamed the National University of Education (Kim et al., 2009b).

Kyungseong Normal College, which was reformed from *Hansung Normal School* in 1895, began secondary school teacher training in 1945. In 1949, it became the College of Education of Seoul National University, while at the provincial level, *Daegu Normal College* and *Gongju Normal College* became public suppliers of secondary school teacher training institutes. Private universities, on the other hand, have been involved in the training of teachers since 1951. *Ewha Woman's University* first opened teacher training programs as a private teacher training institute. When Korea saw a large expansion of secondary education, which resulted in a scarcity of secondary school teachers, numerous private providers started teacher training programs in 1965 to diversify and extend secondary teacher training. However, since the middle of the 1990s, the proliferation of teacher training institutes has resulted in a high level of competition in the current recruitment examination system.

Teacher Training Scale and Current Situation of Teacher Recruiting Thirteen universities offer elementary school teacher education. Secondary education is provided through 46 colleges of education, which include 14 departments of education in general universities, 152 teaching courses in general colleges, and 112 teaching courses in the graduate school of education (Kim et al., 2008: 21–69). Prospective teachers must pass the recruiting examination to work as a teacher at a school (Park et al., 2015: 45–46). The number of graduates from the 2015 elementary school teacher recruitment exam was 4357. However, a total of 9132 persons applied for the exam, and 6173 passed, with a passing rate of 67.6%. From the statistical data retrieved from KEDI Statistics (<https://kess.kedi.re.kr/index>), it was found that the total supply of secondary school teacher certificate holders was 50,828, although only 4.0% of them passed the exam in 2011. In 2015, the government attempted to limit supply by lowering the recruitment rate, which resulted in an 11.6% pass rate. However, the supply-demand gap is far too large to close.

3.2 *In-Service Training and Supervision to Improve Teaching Skills*

The quality of teachers' expertise, which leads to the quality of education, has been systematically monitored in the classroom setting in Korea, labelled as supervision. Supervision is defined as a professional activity that assists teachers in improving their teaching quality and skill. In a restricted sense, it is sometimes defined as educational administration. According to Lee (1984), this perspective of supervision as

offering direction to instructors has prevailed in Korea since the commencement of the new educational system throughout the nation-building period after 1948. Teachers were able to attain their educational goals in the front-line education area because of supervision activity in school, which allowed them to continue their educational research and improve their professionalism. In South Korea, supervision is divided into two categories: in-school supervision and external-school supervision. In-school supervision refers to activities conducted within the school under the leadership of the principal. External-school supervision refers to activities conducted under the supervision of the Office of Education and the Ministry of Education, which are higher levels of education authority than the school.

However, in recent years, supervision has not been of great assistance in improving classes, and it has faced criticism, primarily from higher offices of education and even from school principals, for its bureaucratic control. Traditional supervision may be phased out in favor of new approaches such as consulting, coaching, and mentoring. Nonetheless, in the history of Korean education, the function of supervision in fostering teacher professional growth cannot be overlooked.

3.2.1 In-School Supervision

Preparation of Lesson Plans The planning and execution of lesson plans are at the heart of on-site supervision operations. After the legalization of the National Teachers' Union in 1999, lesson planning became obsolete as a result of labor union collective bargaining. Before 1999, teachers were required to submit lesson plans one week ahead of time and gain the principal's approval. In reality, preparing lesson plans for each class was onerous, and teachers found it difficult to implement lessons in the classroom as intended. Preparing lesson plans and developing teaching materials in this manner was obviously a huge undertaking. The following is the account of a former elementary school teacher from the 1960s.

How would I have written those lesson plans if I had to do it all on my own? There were more than ten class groupings in each grade at the time. After that, each group teacher is responsible for one subject. Group 1 will study the Korean language, group 2 will study mathematics, and group 3 will study music..... When it came to Friday, I just gathered all of the lesson plans from other teachers to copy and edit them for my own usage. Even if you merely duplicate it, you will learn from it, and it may be used to create your own lesson plan. And the grade group leader is in charge of approval before leaving work on Friday, and it goes to the vice-principal and then to the principal for approval on Saturday (when it was not yet a five-day system). I was quite occupied. But, hey, I did it every year, so it was worthwhile, and it became a teacher's habit after that. (Sung**, 70 years old, a teacher and an elementary school principal, and a former superintendent of a school district)

When school education in South Korea began shortly after national independence, it is unclear when the culture and tradition of preparing lesson plans originated. However, it is apparent that it started long ago and from the beginning, and even for seasoned teachers, making lesson plans and preparing for lessons was never easy. Since the year 2000, teachers no longer develop thorough lesson plans for every

class. It relieved instructors of some of their responsibilities, but it also meant that teachers would miss out on opportunities to learn from more experienced teachers. However, some events require teachers to write lesson plans: teachers must prepare a lesson plan once or twice a year for the event of class opening or research class. Lesson plans are also necessary to enter several teaching competitions. Naturally, it is still required in teacher education colleges to teach how to design a lesson plan. Preparing lesson plans bolstered teachers' basic value of teaching ability in a variety of ways.

Open Class and Research Class The open class and research class are two more on-campus monitoring activities. The specifics of how this policy of open class and research class is implemented vary by school, but every year, all schools should plan an open class day with parents, school district supervisors, and fellow instructors. Research workshops are also open to the public during this event, allowing teachers to exchange novel and effective instructional strategies with their peers. Although not all teachers are asked to conduct an open class, every teacher should have one at least once throughout his or her career. There will inevitably be criticisms of the open class, such as that it is only for show and not for actual teaching and learning. However, a teacher's ability to instruct can indeed be enhanced through constructive criticism, allowing the teacher to develop their skills.

Teachers' Learning Community and Group Meetings The teachers' group meeting is the final item on the list of in-school supervisory activities. Teachers' group meetings are recommended to be held once a week and are organized by grades and subjects; for example, teachers of 3rd grade will have a meeting at elementary schools, while teachers at secondary schools will have a meeting organized by subjects. The agenda for the teachers' group meeting is usually for teaching techniques and some issues for worthwhile experiences, as well as preparation for research classes. One of the most essential agenda items may have been how to create test items and score the academic evaluation of formative and summative tests during the semester. The findings of the formative evaluations conducted every month within the context of the standardized national curriculum and textbook system became a significant instrument for students' learning management, while also acting as an independent tool to ask for teacher responsibility. Parents' primary concern is the test results, hence they are extremely sensitive to test outcomes. As a result, the reliability and validity of test items among teachers in the same topic and grade were crucial, and they had to take the form of collaboration to retain fairness as high as feasible. Since the establishment of the KTU (Korea Teachers Union) in 1999, all types of paper-delivered evaluations in schools have been severely limited or abolished under the guises of "procrustean or uniformed exam" and "competitive learning," and the core agenda of student evaluation has gradually vanished, and the teacher's group meeting has lost its vibrancy. In any case, the collaborative culture of teachers' group meetings has made a substantial contribution to the Korean teaching community's professionalism.

3.2.2 External-School Supervision

External-school supervision refers to all related activities and programs carried out by higher supervisory entities such as the Office of Education and the Ministry of Education, which are governed by national laws and systems. Every year, the Minister of Education and the Superintendent of Education set supervision standards to give schools direction and concentration while also providing the required support. Although standards for educational activities have already been set through a uniform national curriculum and textbooks, higher authorities can introduce unique educational activities and propagate new ideas for instructional methods if new educational demands appear in the country.

The Ministry of Education and the Office of Education used the research school system to conduct an experiment in the field and promote it countrywide in order to fulfill particular educational activities (policies) and share new ideas. In addition, when new textbooks are released to correspond to the amended national curriculum, a research school system is implemented as a pilot program before the new textbook is distributed for national usage. In recent years, such actions of higher-level authorities' oversight have been replaced by a variety of educational projects. This transformation, however, faced criticism from school teachers that those projects hinder the development of teaching expertise with autonomy.

3.3 *Standardized National Curriculum*

Korean education is based on a nationally regulated curriculum framework that is changed every seven years. The first curriculum was created in 1954, and since the seventh curriculum was created in 2015, it has been decided to change the curriculum in parts rather than to complete an entire revision. The entire revision of the national curriculum necessitates a lengthy and difficult process to reach consensus among stakeholders, which results in arguments and divides among professional education groups, and, more crucially, the full revision is unable to meet educational demands quickly. The new national curriculum for 2022, on the other hand, is on its way.

Additionally, standardized textbooks and teacher guidebooks based on the national curriculum are released. Those textbooks must pass the ministry of education's rigorous evaluation process. Only a few textbooks are chosen, and along with the physical textbooks, digital textbooks are offered. The national curriculum was used to create a nationally standardized academic evaluation and test. In Korea, a standardized education system might serve as a guideline for teacher quality, with the national curriculum serving as the foundation for instructors to create their own educational abilities.

3.4 *Social and Economic Status for Teachers*

As government employees, teachers are promised a lifetime career with social standing and secure income incentives. The wage system for teachers was not attractive enough to recruit outstanding young people during the economic development phase in the 1960s and 1970s, but their economic compensation was gradually enhanced by the government's persistent efforts. The quality of education in Korea has always been a top priority for Korean parents, who have exerted pressure on the government to maintain it. According to an OECD survey (Table 14.2), elementary teachers with 15 years of experience in Korean national and public schools earn up to \$10,000 more per year than the OECD average. A novice teacher's annual compensation is slightly lower than the average, but it rises as the number of teaching years grows. Teachers' salaries in Korea have the highest purchasing power in the world.

3.5 *Unique Personnel Administration System*

If they stay in the profession until retirement, South Korean public school teachers follow a more or less similar career path: teachers are required to take the role of each homeroom teacher besides the subject teacher, to teach at the assigned schools by rotation, and to work hard enough for promotion to become a school principal.

Homeroom Teacher System The homeroom teacher system allowed for schooling with distinct Korean characteristics. In many OECD countries, middle and high schools lack a classroom teacher system. Instead, students are taught by a classroom teacher and go from one classroom to the next to take classes. All students in South Korea have a homeroom teacher, and these homeroom instructors take the role of parents while students stay at school, to serve as a mentor. From the time children start elementary school until they graduate from high school, the system oversees not only their academic progress but also their whole development, providing counseling and assistance in all aspects of their lives. It is undeniable that homeroom

Table 14.2 Comparison of Korean teachers' salary level with OECD average(as of 2020)

(Unit: USD \$)						
Sorted	Starting teacher			15 years experienced teacher		
	Primary school	Middle school	High school	Primary school	Middle school	High school
Korea	33,477	33,539	32,800	59,103	59,165	58,426
OECD average	34,942	36,116	37,811	48,025	49,701	51,917

Source: OECD (2021). Education at a Glance: OECD Indicators: Table D3.1. Teachers' statutory salaries, based on the most prevalent qualifications at different points in teachers' careers (2020). Retrieved from https://www.oecd-ilibrary.org/education/education-at-a-glance-2021_b35a14e5-en

teachers played a significant part in Korea's educational development. Of course, there are many complaints and avoidances of fatigue directed at homeroom instructors these days, but Korea's educational development has been successful due to this homeroom teacher system.

School Rotation System In Korea, teachers in public schools rotate from one school to another after working at one school for a certain period of time (minimum 3 years and maximum 7 years). The goal of this system is to provide equitable teaching services in remote locations, with good promotion points and monetary recompense for those teachers who choose to serve in these places. Except for Korea and Japan, most countries, especially those with a strong heritage of educational autonomy, lack this structure. During the 1970s industrialization period, this rotation system was strengthened. The system was used to bridge the educational divide between areas by transferring teachers from favored to non-preferred regions, and teachers met pupils from various backgrounds and used the opportunity to try various teaching styles. Students would be able to meet a variety of teachers and obtain a high-quality education regardless of where they live or their socioeconomic status.

The rotation method used to attract young and ambitious instructors by offering incentives, but now the rotation to remote locations is not paid as it once was, resulting in a shortage of teachers with good teaching skills in rural places. Teaching, even in such isolated and impoverished schools, was once a sort of opportunity to take when the economy was not as favorable as it is now. The government also provided incentives in the form of housing, extra allowances, and, most importantly, a system of bonus credits for advancement and transfer to better institutions. However, from a broader viewpoint, good teachers are providing excellent learning circumstances in remote places, and this has helped to improve education quality by reducing the uneven distribution of quality teachers.

Promotion System Teacher and principal are not the same things in many countries: a teacher is someone who teaches in a classroom, while a principal is someone who handles administrative problems in general. The responsibilities of classroom teachers and principals are vastly different. As a result, if a teacher gets promoted to principal, he or she perceives himself or herself to be in a separate position. Teachers' primary responsibility is to instruct pupils, but principals' responsibilities are entirely different: principals are responsible for administering and managing the school. In Korea, however, teachers and principals are all designated as 'The Teacher,' and becoming a principal is a concept of promotion in which the principal is expected to be chosen as the school's best-performing teacher. The most important position in the school is that of a teacher, hence the head of the school should be a teacher. When a teacher becomes a principal, he or she advances to a higher position on the continuum of responsibilities, which is founded on the assumption that the responsibilities are continuous. After working as a teacher in the classroom, they can advance to vice-principal, principal, and scholarship/research posts.

4 Conclusion

Korean teachers are well-known for their high quality, and their social treatment is equally world-class. The high performance of students attests to the quality of the teachers. Students perform well in PISA, TIMSS, and other international comparative studies, but their life satisfaction is at an all-time low. Several international comparison studies, however, demonstrate that teachers and pupils are dissatisfied. This is without a doubt a dilemma. This contradiction was investigated in this chapter through two parts of the teaching profession in South Korea: one through the scientific approach of assessing teaching quality with the ICALT instrument, and the other through a comprehensive review of teacher education/training and policy. The two components were balanced with each other and led to the conclusion that high-quality teaching in South Korea must be the result and outcome of a well-organized teacher system from the beginning of training and recruiting to the end of teachers' well-being as a professional job.

According to the ICALT application for Korean school teachers, they perform at the highest level of teaching quality among the countries involved in the project. This teacher quality is regarded as a highly important factor in Korean students' high performance. Each component of teacher-related policy and implementation was discovered to have served as a driving force in empowering teachers in Korea. Many established traditions of teacher education and policy have contributed to the preservation of high levels of teaching quality. Homeroom teachers, principal promotion from classroom teachers, teacher rotation, good pay, and long-term security, lesson planning and open class, teachers group meetings, and so on are only a few examples of best practice.

This teacher power in Korea was obtained via the building of a professional teaching community during the 70 years since the country's independence in 1948, but since the 2000s, the tradition has been compromised in the name of innovation and future transformation. Instead of losing traditional norms too hastily for the sake of the future, we are encouraged to think carefully and explore creative ways to improve the system. However, in order to perceive and assess the reality of teaching quality and encourage teachers in enhancing their skills, a scientific and objective approach to the observation and measurement of teaching quality should be constructed. From both a practical and theoretical viewpoint, the use of ICALT in this context has proven to be the right approach.

References

- André, S., Maulana, R., Helms-Lorenz, M., Telli, S., Chun, S., Fernandez-Garcia, C.-M., de Jager, T., Irmidayanti, Y., Inda Caro, M., Lee, O., Safrina, R., Coetzee, T., & Jeon, M. (2020). Student perceptions in measuring teaching behavior across six countries: A multi-group confirmatory factor analysis approach to measurement invariance. *Frontiers in Psychology, 11*, 273. <https://doi.org/10.3389/fpsyg.2020.00273>

- Chun, S., Lee, O., & Jeon, M. (2017). A study on the analysis of teachers' classroom teaching expertise using ICALT observation tools. *Educational Engineering Research*, 33(2), 517–536.
- Flexner. (1910). *A. Medical education in the United States and Canada* (Bulletin No. 4). Carnegie Foundation for the Advancement of Teaching.
- http://www.nl.go.kr/app/nl/search/common/download.jsp?file_id=FILE-00008153810
- <https://www.moe.go.kr/boardCnts/view.do?boardID=294&lev=0&statusYN=W&s=moe&m=0204&opType=N&boardSeq=79191>
- <https://www.moe.go.kr/boardCnts/view.do?boardID=294&lev=0&statusYN=W&s=moe&m=0204&opType=N&boardSeq=81851>
- <https://www.oecd.org/education/talis/talis-2018-data.htm>
- <https://www.oecd.org/pisa/publications/pisa-2018-snapshots.htm>
- <https://www.rug.nl/gmw/lerarenopleiding/onderzoek/psychometrisch/>
- <https://www1.compareyourcountry.org/talis-teacher-survey/en/10/1836+1838/ranking/>
- Kim, O. (2006). A study on the reconceptualization of teacher professionalism. *The Journal of Educational Administration*, 24(4), 139–160.
- Kim, T., et al. (2008). *A study on the diversification of teacher training and recruitment*. Ministry of Education, Science and Technology.
- Kim, E., Kim, J., & Han, Y. (2009a). *Secondary education and teacher quality in the Republic of Korea* (A Series Book of Secondary Teacher Policy in Asia). Unesco.
- Kim, G., Park, Y., Jung, G., Kim, K., & Kim, B. (2009b). *A study on the reform of the teacher training system* (RR2009-02). Korea Educational Research Institute.
- Lee, J. (1984). Notes on the concept and system of scholarship. *The Journal of Educational Administration*, 2, 19–43.
- Lieberman, M. (1956). *Education as a profession*. Prentice-Hall, Inc.
- Marzano, R. J., Pickering, D., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. ASCD.
- Maulana, R., Smale-Jacobse, A., Helms-Lorenz, M., et al. (2020a). Measuring differentiated instruction in The Netherlands and South Korea: Factor structure equivalence, correlates, and complexity level. *European Journal of Psychology of Education*, 35, 881–909. <https://doi.org/10.1007/s10212-019-00446-4>
- Maulana, R., André, S., Helms-Lorenz, M., Ko, J., Chun, S., Shahzad, A., et al. (2020b). Observed teaching behavior in secondary education across six countries: Measurement invariance and indication of cross-national variations. *School Effectiveness and School Improvement*, 1–32.
- Maulana, R., André, S., Helms-Lorenz, M., Ko, J., Chun, S., Shahzad, A., Irmidayanti, Y., Lee, O., de Jager, T., Coetzee, T., & Fadhilah, N. (2020c). Observed teaching behaviour in secondary education across six countries: Measurement invariance and indication of cross-national variations. *School Effectiveness and School Improvement*, 32, 64. <https://doi.org/10.1080/09243453.2020.1777170>
- OECD. (2019). *TALIS 2018 results: Teachers and school leaders as lifelong learners*.
- OECD. (2021). *Education at a glance: OECD indicators*.
- Park, S., Park, S., Lee, I., Lee, G., Park, Y., Maeng, J., & Woo, H. (2015). *A study on the reorganization of the teacher training system*. Ministry of Education Policy Research Report.
- van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research*, 49(2), 127–152. <https://doi.org/10.1080/00131880701369651>
- van de Grift, W. J., Chun, S., Maulana, R., Lee, O., & Helms-Lorenz, M. (2017). Measuring teaching quality and student engagement in South Korea and The Netherlands. *School Effectiveness and School Improvement*, 28(3), 337–349.
- van de Grift, W., Lee, O., Maulana, R., Chun, S.-Y., & Helms-Lorenz, M. (2019). Measuring teaching skill of South Korean teachers in secondary education: Detecting a teacher's potential zone of proximal development using the Rasch model. *Studies in Educational Evaluation* (forthcoming online publication first).

Prof. Seyeoung Chun is Professor Emeritus of Education at Chungnam National University, one of the major national universities in Daejeon, Korea. He received his education and Ph.D. from Seoul National University, South Korea, and has been actively engaged in education policy research and has held several key positions such as Secretary of Education to the President and CEO of KERIS. He founded the Smart Education Society in 2013, and has led many projects and initiatives for the paradigm shift of education in the digital era. Since his early career at the Korean National Commission for UNESCO, he has participated in many international cooperation projects and worked for several developing countries such as Nicaragua, Honduras, Cambodia, etc. Education Miracle in the Republic of Korea is the latest book to be published as a summary of his academic life.

Prof. Okhwa Lee is a specialist in educational technology and a practitioner of teacher education. She has been a pioneer of the e-learning, technology applications in education and educational reform through smart education in Korea. She was a member of the Presidential Educational Reform Committee and the Presidential e-Government Committee of the Republic of Korea, also consulting members for various ministries regarding educational applications of technology. She has rich experiences of international collaborations with the Europe Erasmus mobility with Finland Sweden, Estonia, Netherlands and etc., long history of research collaboration with USA, Australia, Thailand and etc. Recently she collaborated with developing countries through the Korean government ODA (Official Development Assistant) programs to Sudan, Nigeria, Nicaragua, Vietnam, Ethiopia, Cambodia, Myanmar, and etc. Her work through ODA focused on teachers' capacity development of teaching skills using technology.

Dr. Deuk-Joon Kim has been active in the fields of 'Electronics Engineering' and 'Information Security' from the late 1980s to the early 2000s. Since then, I've been pursuing education (majoring in 'Educational Technology') and have obtained a Master's degree (2006) and a Doctorate degree (2014) in education. I have worked as a professor at private and national universities and have conducted various research in the field of educational technology, such as development and utilization of educational technology and educational evaluation. Recently, I've been working as an ODA expert in the field of education, participating in various domestic and international education projects and research, and collaborating with domestic and foreign institutions to promote educational innovation and development. In addition, I have presented excellent research outcomes in the field of educational technology, such as publishing papers in numerous international academic journals and presenting at domestic and international academic conferences.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 15

Classroom Learning Environments and Assessment Practices in Science Classrooms in Western Australia



Rekha B. Koul

Abstract The research described in this paper was aimed at identifying exemplary assessment practices in secondary science classes. In the first stage, following a review of the literature, a six-scale instrument of 48 items was trialed with a sample of 470 students from grades eight, nine and ten in 20 science classrooms in three Western Australian schools. Based on internal consistency reliability data and exploratory factor analysis, refinement decisions resulted in a five-scale instrument that was named the *Student Perceptions of Assessment Questionnaire* (SPAQ). In the second stage, the SPAQ was used with an attitude scale, and a self-efficacy scale. This survey was administered to a larger sample of 960 students from 40 science classes from the same grades as in the first stage. Statistical analyses confirmed the validity and reliability of the SPAQ. Based on the results of this survey exemplary teachers were identified. In the third and last stage interviews with teachers and students were conducted. Classes of these exemplary teachers were also observed. These exemplary teachers were found to be thorough in their teaching, giving students enough time to prepare for the assessment, giving students freedom to choose from a variety of assessments and were flexible in teaching and assessment. They also demonstrated in-depth understanding of the science topics they were teaching.

Keywords Assessment practices · Student perceptual data · Exemplary teachers

1 Introduction

A constructivist view of learning supports the use of clear goal statements and success criteria, targeted feedback and student self-assessment (Muijs & Reynolds, 2017, p. 1; Sadler, 1989). This idea is in line with effective teaching research (Maulana et al., 2021). However, little contemporary evidence exists to support the view that students are genuinely involved in decision-making about their

R. B. Koul (✉)

STEM Research Group, School of Education, Curtin University, Perth, WA, Australia
e-mail: R.Koul@curtin.edu.au

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_15

317

assessment tasks (Dorman et al., 2008). That is, forms of assessment and specific assessment tasks employed in schools are usually decided by teachers and administrators. Furthermore, even though reports like *The Status and Quality of Teaching and Learning in Australia* (Goodrum et al., 2001) have asserted that assessment is a key component of the teaching and learning process, teachers tend to utilize a very narrow range of assessment strategies on which to base feedback to parents and students. In practice, there is little evidence that teachers actually use diagnostic or formative assessment strategies to inform planning and teaching (Radnor, 1996).

There are conflicting views about the role and nature of assessment practices in education. Harlen (1998) advocates that teacher should use both oral and written questions in assessing student's learning. While, experts (Dorr-Bremme & Herman, 1986; Stiggins, 1994) encourage alternative assessment strategies, such as teacher observation, personal communication, and student performances, demonstrations, and portfolios, for greater usefulness of evaluating students and informing classroom instruction. Tobin (1998) asserted that assessment can be used to provide opportunities for students to show what they know. Reynolds et al. (1995) argued that for effective learning to occur, congruence must exist between instruction, assessment and outcomes. This paper represents a context-specific investigation of this congruence.

An effective assessment process should involve a two-way communication system between teachers and their students (Black & William, 1998). Historically, teachers have used testing instruments to transmit to the student and their parents what is really important for the student to know and do. While this reporting tends to be in the form of a grade, the form and design of the assessment can send subtle messages on what is important. There has been a substantial amount of research into types of assessment but very little research into students' perceptions of assessment (Black & William, 1998; Crooks, 1988; Plake, 1993; Popham, 1997) and how it relates to classroom learning environments.

2 Aim

The overall aim of the study was to investigate relationships among students' perceptions of their assessment tasks, classroom learning environments, academic efficacy and attitude to science in years eight, nine and ten in Western Australia.

The objectives of this study were:

1. to provide further validation data on the instrument for accessing students perceptions of assessment tasks;
2. to investigate differences between students' perceptions in terms of gender and year levels;
3. to investigate associations between students' perceptions of their assessment tasks and their attitude to science and academic efficacy outcomes; and
4. to describe the form and design of assessment tasks used by exemplary science teachers.

3 Theoretical Framing

3.1 *Use of Student Perceptual Data*

Until the late 1960s a very strong tradition of trained observers coding teacher and student behaviors dominated classroom research. Indeed, it was a key recommendation of Dunkin and Biddle (1974) that instruments for research on teaching processes, where possible, should deal with the objective characteristics of classroom events. Clearly, this approach to research which often involved trained observers coding teacher and student behaviours was consistent with the behaviourism approach of the 1960s. The study of classroom psychosocial environments in the late 1960s broke this tradition and used student perceptual data. Since then, the strong trend in classroom environment research has been towards this high-inference approach with data collected from the teachers and students. Walberg (1976) supported this methodological approach where student learning involves student perceptions acting as mediators in the learning process. Walberg (1976) also advocated the use of student perception to assess learning environments because students seemed quite able to perceive and weigh stimuli and to render predictively valid judgments of the social environments of their classes.

3.2 *Classroom Learning Environment*

The notion that a learning environment exists which mediates aspects of educational development began as early as 1936 when Lewin (1936) recognised that the environment and the personality of the individual were powerful determinants of behaviour and introduced the formula, $B = f(P,E)$. Since Lewin's time, international research efforts involving the conceptualisation, assessment, and investigation of perceptions of aspects of the classroom environment have firmly established classroom environments as a thriving field of study (Fraser, 1994, 1998; Fraser & Wallberg, 1991). For example, classroom environment research has focused on constructivist classroom environments (Taylor et al., 1997), cross-national constructivist classroom environments (Aldridge et al., 1999), science laboratory classroom environments (McRobbie & Fraser, 1993), computer laboratory classroom environment (Newby & Fisher, 1997) computer-assisted instruction classrooms (Stolarchuk & Fisher, 1999) and classroom environment and teachers' cultural back grounds (Koul & Fisher, 2006).

A great deal of classroom learning environment research has been carried out over the past 40 years and evidence from these studies reveals that classroom learning environment dimensions are good indicators of teaching and learning processes and have predictive power on a number of learning outcomes pointing towards the possibility of improving students' outcomes through changing classroom environments (Fraser, 1994, 1998; Fraser & Wallberg, 1991; Wubbles & Levy, 1993). The

present interpretive study involved a multi-method approach in exploration of factors associated with students' perceptions of assessment.

3.2.1 Attitude to Science Classrooms

The impact of students' attitudes towards their science assessments is regarded as an important goal in the present study. Attitudes towards science, has been defined as "a learned disposition to evaluate in certain ways objects, people, actions, situations or propositions involved in learning science" (Gardner, 1975, p. 2). This learned disposition refers to the way students regard science, such as interesting, boring, dull or exciting. Positive student attitudes are then measured by the degree of motivation and interest reported by the students. Klopfer (1971, 1976) went further and developed a structure for evaluating attitudes related to science education. He included four categories in his structure: events in the natural world; activities; science; and inquiry. Klopfer's (1976) second category, relating to students attitudes towards their science assessments was a focus of the present study.

3.2.2 Academic Efficacy

Over the past two decades the broad psychological concept of self-efficacy has been a subject of interest (Bandura, 1997; Schunk, 1995). Within this field, one particular strong area of interest is that of academic efficacy, which refers to personal judgments of one's capabilities to organize and execute courses of action to attain designated types of educational performances (Zimmerman, 1995). Research studies have provided consistent, convincing evidence that academic efficacy is positively related to academic motivation (e.g., Schunk & Hanson, 1985), persistence (Lyman et al., 1984), memory performance (Berry, 1987), and academic performance (Schunk, 1989).

3.2.3 Gender and Year Level

It is well-documented in reviews of literature that women are under-represented in science and technology courses and careers (Commonwealth of Australia, 2019; Greenfield, 1996; Kahle & Meece, 1994) and that boys outperformed girls in science (especially physical science) (Casad et al., 2018; Bellar & Gafni, 1996; Kahle & Meece, 1994; Murphy, 1996). Among the sources that may cause these differences are individual, cognitive, attitudinal, socio-cultural, home and family, and educational variables (Farenga & Joyce, 1997; Kahle & Meece, 1994). In the classroom context, boys and girls may not have equal opportunities in science activities, and this could cause gender differences in science achievement (Fraser et al., 1992; Harding, 1996; Warrington & Younger, 1996). Because educational variables are one of the important sources for accounting for gender differences in students'

achievement in science, and for participation in science activities, the perspective of gender differences needs to be understood. Previous studies have reported gender-related differences in students' perceptions of the learning environment (Fraser et al., 1996; Koul & Fisher, 2006). Therefore in keeping with these lines of research, gender-related differences in students' perceptions of their assessment were explored in this study.

Year level as well as gender differences in students' perceptions, other learning environment research studies in science classrooms have indicated differences between perceptions of students in different years of school (Kim et al., 2000). In this study, differences between the perceptions of students in different years of lower secondary were examined for trends.

4 Instruments and Procedure Used

The study was carried out in phases over a period of three years using a multi-method research approach:

1. In the first phase a pre-existing and validated questionnaire, *Perceptions of Assessment Tasks (PAT)* a six-scale instrument of 55 item developed by Schaffner et al. (2000) was administered to 470 students from grades eight, nine and ten in 20 science classrooms in three Western Australian schools. Students in this study were between the ages of 12–15 years. Close ended interviews were conducted with randomly selected 40 students to look at student perceptions of their assessment tasks.
2. In second phase based on internal consistency reliability data and exploratory factor analysis, refinement decisions of PATT resulted in a five-scale instrument that was named the *Student Perceptions of Assessment Questionnaire (SPAQ)*. This study was part of a larger study carried out in three states of Australia. The SPAQ was used with an attitude scale, and a self-efficacy scale. This survey was administered to a larger sample of 960 students from 41 science classes from the same grades as in the first stage.
3. In the final stage of the study five teachers identified on the basis of students showing most positive perceptions on the scales of SPAQ were interviewed and their teaching observed. Informal interviews were also conducted with students from the classes identified.

Students' Perceptions of Assessment Questionnaire (SPAQ) Students' perceptions of assessment were assessed with the 30-item SPAQ. These items are assigned to internally consistent scales namely Congruence & Planned Activity, Authenticity, Student Consultation, Transparency and Diversity. Table 15.1 shows the scales, descriptions and sample items from the SPAQ. Validation statistics performed on the data collected are presented in the results section. Responses in the SPAQ were recorded on a four point Likert type response format for each item (e.g., Almost Never, Sometimes, Often, and Almost Always).

Table 15.1 Description and example of items for each Scale of Students Perceptions of Assessment Questionnaire (SPAQ), attitude scale and academic efficacy

Scale	Description	Item
Congruence and Planned Activity	Extent to which assessment tasks align with the goals, objectives and activities of the learning program.	My assignments/tests are about what I have done in class.
Authenticity	The extent to which assessment tasks feature real life situations those are relevant to the learner.	I find science assessment tasks are relevant to what I do outside of school.
Student Consultation	The extent to which students are consulted and informed about the forms of assessment tasks being employed.	I have a say in how I will be assessed in science.
Transparency	The extent to which the purposes and forms of assessment tasks are well-defined and clear to the learner.	I am clear about what my teacher wants in my assessment tasks.
Diversity	The extent to which all students have an equal chance at completing assessment tasks.	I have as much chance as any other student at completing assessment tasks.
Attitude to Science	The extent to which students are interested in, enjoy and look forward to lessons in that subject.	I enjoy the activities we do in science.
Academic Efficacy	Students' judgments of their capabilities to organize and execute courses of action to attain designated types of educational performances.	Even if science is hard, I can learn it.

Two outcome scales namely Attitude to Science and Academic Efficacy were also employed in present study. A review of literature revealed a large pool of science-related attitude scales. Of particular interest to this study is the *Test of Science Related Attitudes* (TOSRA) developed by Fraser (1978) to measure students' attitudes towards their science classes. Fraser based the subscales of this instrument on Klopfer's (Klopfer, 1976) taxonomy of the affective domain related to science education. Attitude to Science was assessed on a 8-item scale adopted from the *Test of Science-Related Attitudes* (TOSRA: Fraser, 1981). Responses were recorded on a four-point format ranging from 1 (Disagree) to 4 (Agree).

Perceived Academic Efficacy refers to students' judgments of their ability to master academic tasks that they are given in their classrooms. A 6-item scale using items developed by Midgley and Urdan (1995) was used to assess perceived academic competence at science class work. Items were modified to elicit a response on academic efficacy in science. All items in the academic efficacy scale had a four-point response format with anchors of 1 (Disagree) and 4 (Agree).

5 Results

Results of the study are presented in lieu of each of the research objectives:

5.1 Objective 1: Validation Data on the Instrument for Accessing Students' Perceptions of Assessment Tasks

A principal components factor analysis followed by varimax rotation confirmed a refined structure of the SPAQ instrument comprising of 30 items in 5 scales and 14 items in two outcome scales. All the 44 items had a loading of at least 0.40 on their *a priori* scales (see Table 15.2). The percentage of the total variance extracted with each factor is also recorded at the bottom of Table 15.2. The percentage of variance varies from 3.55% to 26.03% for different scales, with the total variance accounted for being around 50%.

The validity and reliability information of the instrument developed in this study are presented in Table 15.3.

To determine by the degree to which items in the same scale measure the same aspects of students' perceptions of assessment tasks, attitude to science and academic self-efficacy, a measure of internal consistency, the Cronbach alpha reliability coefficient (Cronbach, 1951) was used. For the scales of SPAQ, the highest alpha reliability of 0.83 for the scale of Authenticity, and the lowest of 0.63 for the scale of Diversity was recorded. The scale of student attitudes to science has alpha reliability score of 0.85 and scale of Academic Efficacy of 0.90. Since all the reliabilities for the scales of SPAQ were consistently above 0.63 the instrument developed is therefore reliable for use (DeVellis, 1991).

High mean scores ranging from 2.16 for the scale of Student Consultation to 3.17 for the scale of Congruence with Planned Learning on a four-point Likert type scale confirm that students generally have a positive perception of their assessment tasks. Scale of Student Consultation having the lowest scores confirms that students generally do not have a say in their assessment tasks.

Overall culture of each class is different and the ability of SPAQ to differentiate between the classes in the study was considered important. The instruments' ability to differentiate in this way was measured using one-way analysis of variance (ANOVA). The eta² statistics was calculated to provide an estimate of the strength of the association between class membership and the dependent variables as shown in Table 15.3. The eta² statistic for the SPAQ, indicates that the amount of variance in scores accounted for by class membership ranged from 0.12 to 0.28 and was statistically significant ($p < 0.001$) for all scales. It appears that the instrument is able to differentiate clearly between the perceptions of students in different classrooms.

Table 15.2 Factor loadings for the questionnaire used in the study

Item no	Congruence and planned activity	Authenticity	Student consultation	Transparency	Diversity	Attitude to science	Academic efficacy
1	0.46						
2	0.52						
3	0.51						
4	0.68						
5	0.69						
6	0.46						
7		0.43					
8		0.60					
9		0.69					
10		0.64					
11		0.72					
12		0.52					
13			0.47				
14			0.49				
15			0.66				
16			0.59				
17			0.44				
18			0.74				
19				0.66			
20				0.63			
21				0.68			
22				0.69			
23				0.64			
24				0.57			
25					0.41		
26					0.48		
27					0.54		
28					0.44		
29					0.46		
30					0.52		
31						0.77	
32						0.47	
33						0.66	
34						0.71	
35						0.44	
36						0.78	
37						0.60	
38							0.42
39							0.61
40							0.73
41							0.69

(continued)

Table 15.2 (continued)

Item no	Congruence and planned activity	Authenticity	Student consultation	Transparency	Diversity	Attitude to science	Academic efficacy
42							0.77
43							0.76
44							0.71
% Variance	26.03	8.12	7.05	4.13	3.55	3.43	2.46
Eigen value	11.45	3.57	3.10	1.81	1.56	1.50	1.08

Table 15.3 Scale Mean, Standard Deviation, Internal Consistency (Cronbach Alpha Reliability) and ability to differentiate between classrooms (ANOVA Results) for the SPAQ, attitude to science and academic efficacy

Scale	Mean	St. Dev	Alpha reliability	ANOVA (η^2)
Congruence and planned activity	3.17	0.51	0.76	0.13*
Authenticity	2.24	0.64	0.83	0.17*
Student consultation	2.16	0.58	0.71	0.19*
Transparency	3.06	0.63	0.83	0.14*
Diversity	2.56	0.54	0.63	0.12*
Attitude to science	2.56	0.80	0.85	0.28*
Academic efficacy	2.96	0.76	0.9	0.13*

n = 960 students in 40 classes * $p < 0.001$

5.2 Objective 2: Differences Between Students' Perceptions in Terms of Gender and Year Levels

5.2.1 Gender Differences

Differences between the students' perceptions of the scales of the SPAQ and the gender of the students were analysed. The gender differences in students' perceptions of classroom learning environment were examined by splitting the total number into female (388) and male (572) students involved in the study.

To examine the gender differences in students' perceptions of the classes, the within-class gender subgroup mean was chosen as the unit of analysis as this aims to eliminate the effect of class differences due to males and females being unevenly distributed in the sample. In the data analysis, male and female students' mean scores for each class were computed, and the significance of gender differences were analysed using an independent t-test. Table 15.4 shows the scale item means, male and female differences, standard deviations, t-values and Cohen's d effect size. The purpose of this analysis was to establish whether there are significant differences in perceptions of students according to their gender.

As can be seen in Table 15.4, out of five scales of the SPAQ and two Attitude scales, the gender differences in the perceptions of males and females were found to

Table 15.4 Item mean and standard deviation for gender differences in students' perceptions on the scales of SPAQ

Scale	Item mean		Item SD		Difference	
	Female	Male	Female	Male	t	Effect Size
Congruence with planned activity	3.21	3.14	0.51	0.51	0.05	0.13
Authenticity	2.18	2.28	0.67	0.61	5.28*	0.15
Student consultation	2.12	2.19	0.58	0.58	0.06	0.12
Transparency	3.07	3.06	0.62	0.65	0.52	0.01
Diversity	2.54	2.58	0.54	0.54	0.26	0.07
Student attitudes	2.52	2.58	0.81	0.78	0.01	0.07
Academic efficacy	2.88	3.01	0.77	0.74	0.48	0.17

* $p < 0.05$, females ($n = 388$); males ($n = 572$)

Table 15.5 Item Mean, Item Standard Deviation and ability to differentiate between levels (ANOVA results) for year level differences in students' perceptions measured by the SPAQ

Scale	Mean			SD			Difference
	Year 8	Year 9	Year 10	Year 8	Year 9	Year 10	F
Congruence with planned activity	3.24	3.14	3.12	0.51	0.49	0.51	5.64**
Authenticity	2.32	2.24	2.15	0.66	0.65	0.59	5.13**
Student consultation	2.34	2.15	1.95	0.56	0.59	0.51	37.06***
Transparency	3.10	3.09	2.99	0.64	0.64	0.62	2.48
Diversity	2.62	2.58	2.47	0.54	0.55	0.53	6.21**
Student attitudes	2.65	2.44	2.60	0.81	0.85	0.71	6.38**
Academic efficacy	2.99	2.96	2.91	0.74	0.76	0.78	0.83

Sample Size = 347 (Year 8), 328 (Year 9) and 285 (Year 10)

** $p < 0.01$, * $p < 0.001$, *** $p < 0.05$

be statistically significantly different only on the scale of Authenticity. The result indicates that Authenticity was reported higher by male compared to female students.

5.2.2 Year Level Differences

One of the aims of the study was to investigate the differences in the perceptions of the scales of SPAQ and the two sides of attitude and efficacy in students from different year levels. This was explored by splitting the students in their year groups (year 8 = 347, year 9 = 328, year 10 = 285).

The results of the analyses are shown in Table 15.6. In the data analysis, mean scores for each of the three-year groups were computed. Table 15.5 shows the scale item means and F values of the scales of the SPAQ with the perceptions of students from the three year groups in study. The purpose of this analysis is to establish whether there are Significant differences in the perceptions of students according to their year groups.

Table 15.6 Associations between scales of SPAQ and attitude to science in terms of simple correlations (R), multiple correlations and standardized regression coefficient (β)

Scale	Attitude to science class		Academic efficacy	
	r	β	R	β
Congruence and planned activity	0.32**	0.10***	0.33**	0.09***
Authenticity	0.45**	0.35***	0.36**	0.22***
Student consultation	0.25**	-0.17***	0.21**	-0.16***
Transparency	0.39**	0.18***	0.46**	0.28***
Diversity	0.39**	0.18***	0.43**	0.25***
Multiple correlations	$R = 0.55^*$		0.56^*	
	$R^2 = 0.3$		0.32	

* $p < 0.0001$, ** $p < 0.01$, *** $p < 0.05$ $n = 960$

As can be seen in Table 15.6, the differences in the perceptions of students on the scales of SPAQ and Attitude, five out of seven scales are statistically significant confirming that year level does impact significantly on students' perception of their assessment. Tukey's post hoc test ($p < 0.05$) revealed that for the Congruence with Planned Activity scale the Year Eight students were dominant and had statistically significant higher means while the Year Ten students had the highest means for the scale of Diversity.

5.3 Objective 3: Associations Between SPAQ and Attitude to Science and Academic Efficacy

One of the aims of the study was to investigate associations between students' perceptions of assessment tasks and their attitude to science classes. These associations were explored using simple and multiple correlation analyses. The results of the analyses are shown in Table 15.4. For all the scales of the SPAQ associations are positive and statistically significant.

It was found that the scales of Congruence and Planned Activity, Authenticity, Transparency and Diversity were positively and significantly associated whereas, scale of Student Consultation was negatively and significantly associated with attitude to science.

The multiple correlation (R) between the set of SPAQ scales and attitude to science class was 0.55. The R^2 value which indicates the proportion of variance in attitude to science class that can be attributed to students' perceptions of their assessment tasks given by the teachers was 30%. To determine which SPAQ scales contributed most to this association, the standardized regression coefficient (β) was examined for each scale. It was found that the scales of Congruence and Planned Activity, Authenticity, Transparency and Diversity were positively and significantly associated whereas, scale of Student Consultation was negatively and significantly associated with attitude to science.

5.4 Objective 4: Describe the Form and Design of Assessment Tasks Used by Exemplary Science Teachers

Based on the findings of the quantitative data five exemplary teachers (three male and two female) were identified from the total sample of 40 and their teaching observed and informal interviews conducted. These five teachers represented Private, Public and Rural schools in Western Australia. These selected teachers had been rated by their students' more than one standard deviation above the mean for at least three of the five scales. This process has been described previously by Waldrup et al. (2009).

Furthermore, four students from the classes of each of the five selected teachers also were interviewed. The students' interviews were structured and conducted in three phases on the same day. The interview phases occurred before, during and after an activity in the classroom. Similar questions regarding the activity were asked to assess students' initial perceptions about the task, during the task and when the task was completed.

The students were asked few general questions followed by questions relating to each of the five scales of SPAQ questionnaire. This approach enabled the researcher to draw on a variety of paradigms to inform their interpretation in a bid to explain the positive student perception of assessment tasks. The interview schedule along with stages and scales is represented in Table 15.7.

Table 15.7 Student interview schedule

Scale	Beginning	During	After
General	What do you think of the task? Why do you feel that way? How do you feel you will achieve in this task? Why do you feel that way?	What is challenging about the task? Why do you think that is? What are you learning during this task? Have your feelings towards the task changed? What has changed? What would you like to do differently?	What have you learned during this task? Did you like learning in this way? Why do you think that? Was it challenging? Why do you think that is? Was there anything different about the task? What would you like to have done differently?
Congruence	Is the task related to what you have been learning in class?	Is the task related to what you have been learning in class?	Was the task related to what you have been learning in class?
Authenticity	Is the task useful or helpful to you? Is it meaningful to you?	Is the task useful or helpful to you? Is it meaningful to you?	Was the task useful or helpful to you? Was it meaningful to you?
Student consultation	Do you have some say in this task?	Are you having some say in this task?	Did you have some say in this task?
Transparency	Are you clear about what you need to do in this task?	Are you still clear about what you need to do in this task?	Were you clear about what you needed to do in this task?
Diversity	Are there different ways you can complete this task? What will happen if you are confused?	Are there different ways you can complete this task? What will happen if you are confused?	Were there different ways to complete this task? What happen if you are confused?

The results which emerged from the interviews with teachers and students are presented in the next section.

Learning and Assessment Interviews and observations reflected that the exemplary teachers were engaging constructivist ways of teaching underpinning formulations of formative assessment (Sadler, 1989). As supported by the quantitative results, students of these teachers had very positive perceptions of the assessment practices employed by their teachers and it was observed that social interactions within these classes were generally very strong. Assessment practices employed by these teachers not only look at what students know, but also at developing student identities as capable and competent learners. These teachers take into consideration what, why, and how students are learning as well as showing a shift in their views of assessment in science by keeping themselves informed on the changing nature of the outcomes of the science education. Some of the comments supporting these claims are:

Teacher:	I formulate assessments very early in the year keeping science intended outcomes in view. My assessments are designed to let me know what students know, not what they do not know. Thus, assessment becomes a part of learning.
Student 1:	Beginning of each term he gave us details of all the assessment and what is expected of us. This approach gives us clear guidelines for learning. After the assessment is evaluated, often, he runs a session on our misconceptions.
Student 5:	We knew in the beginning of the term that we are required to make an information poster or pamphlet or flyer regarding the infectious diseases. I kept on collecting the related information and stuff you know... It was easy to compile all the information close to the date of submission
Student 8:	Since we know what is required and even expected from us... we learn accordingly.
Student 11:	If I decide that I want good marks, we have to work for it. I cannot say what the mark should be but, If I have worked according to teachers guidelines I am sure that my work will get a high mark.

Curriculum and Assessment The teachers when interviewed commented on the way they considered assessment and curriculum to be related and interact in complex ways. They believed that a well perceived curriculum that incorporates assessment also narrows the gap between intended and implemented curriculum resulting in an achieved curriculum. Exemplary teachers also researched and used the available relevant assessment resources. Typical of their comments were:

Teacher:	I do not separate assessment from the curriculum. Both are different but lead to same object-student/teacher learning. It is complex but once understood can be practiced successfully. These days there are lot of ideas and materials available.
Student 4:	She tells us what will be asked to do. So we prepare accordingly. She also gives us an evaluation criterion for each assessment.

(continued)

Student 9:	You know this was different. I exactly knew what is required in this project. It turned out to be the biggest project I had ever done.
Student 14:	The last work sheet he gave us was confusing to start with. Lot of application mixing topics in machines, light and heat. I thought about it and did well.

Classroom and Assessment The exemplary teachers believed that there is a need to recognise the roles and responsibilities of both teachers and students. This view resonates with Sadler's (1989) view that formative assessment is based on the principle that students need to become consumers as well as the objects of assessment activities. This sociocultural view of learning enhances positive classroom interactions. Assessments also reflect a power relationship in classroom. The teacher questions and students respond. However, in an exemplary teacher's class, teacher provides enough resources for students to respond to the questions and create knowledge. These resources could be books, the World Wide Web, peers or other resource persons.

Teacher:	I provide many resources to students so that they can research and find answers to the investigations we do. It is interesting to see how many resources students find on their own and enter classroom with different world views.
Students 3:	Teacher directs us to the reading material. We also do lot of web surfing. I find many useful links on YouTube.
Student 7:	Last night when I was chatting with my friends on Facebook [internet interaction site] we looked at viruses, bacteria, protozoa, worms and fungi. That was cool. We all learnt a lot about the lesson we are doing in class.
Student 10:	First, I thought we are not going to learn much in this year's science unit. It seemed he was boring. I had not done much research. Now that we have started researching and we find the importance of substances like the mining in up north. We get the crude material and useful things come out of that.

Teachers and Assessment Although these selected teachers had emancipatory views about assessment and stood apart generally from their counter-parts, they were feeling concerned about the external influences on them. They felt answerable to various stake holders namely students, parents, administrators and the community at large. To establish their accountability their students had to perform well in national and international science tests. They could use these test results as evidence of efficiency for their performance. The teachers also believe that knowledge and expertise of various assessment activities is mandatory for all science teachers who need to have an in-depth understanding of the topic being taught and that students' existing knowledge. The exemplary teachers recommend that this can be achieved through planning of the course content which should include teaching, learning, assessment and curriculum and their interrelationship.

Teacher:	I feel responsible for student learning. I am answerable for the student learning and on top of that we have science Olympiads, national testing and international testing. It is complex.
Student 2:	He knows his stuff well and also how to teach. For example last topic on renewable energy he talked about many ways, how energy can be renewed and also conserved. It was great. I enjoyed the lesson and writing the project. With the result I got good grade.
Student 16:	She is through with the content of all the lessons.
Student 15:	Later this year we will be writing the international science test and she wants us to do well in that. This is a science extension class and many of us also participate in science Olympiads.

Students and Assessment The final and last section of this study identified the students as active and intentional participants in classroom assessment practices. Cowie (2005) highlights the multiple consequences of classroom assessment for students as: importance of trust and respect; the influence of their goals and learning motivations, and equity issues. Our study also found parallels with each of these factors. Continued teacher support and positive classroom learning environment contribute towards what students consider important to learn. Mutual trust and respect among teachers and students is central to student learning. Students should believe that assessments are designed to help them and they view assessment as a joint teacher-pupil responsibility.

Teacher:	I have to be very careful about what I speak in classroom. I try to look at students positive points and build on that. I tend to add plurality in the assessments we (students and I) design. This gives all students from different cultural backgrounds and ability levels to demonstrate their learning. It also keeps them interested in science.
Student 6:	What I love about our teacher is the respect and belief she has for us. She designs assessments which she is confident that we have learnt and can do well. Last assignment when she thought that I could improve upon it, she talked privately and respectfully to me. I am learning, and that is her job.
Student 18:	During the question/answer session every student has equal chance of being asked for a response. He will only ask those students who have raised their hands. In the class (while teaching) he never shows individual preference.
Student 12:	We are free to do our assignments the way we want. We don't get a choice on the things where teacher has already planned an activity and if we change it would affect our learning.

6 Discussion and Conclusion

This study further validated an instrument the Students' Perception of Assessment Questionnaire (SPAQ) for use in educational settings. The three stage data collection facilitated gaining in-depth insights into students perceptions of assessments

and how students felt assessment as an integral part of learning and playing significant role in teacher and student behaviours in the classroom (Cowie, 2005). The questionnaire using student perceptual data (Walberg, 1976) scales showed an acceptable factor loading with 30 items in five scales and Cronbach alpha reliability scores ranged from 0.63 to 0.83, (DeVellis, 1991), thus making these scales acceptable for use in future. Study made use of the student perception of assessment tasks added to the existing paucity of research in this area (Black & William, 1998; Crooks, 1988; Plake, 1993; Popham, 1997).

Of five scales of the questionnaire lowest mean score was recorded for the scale of Student Consultation which confirms that students generally are not consulted when deciding about the types of assessments and are not involved a two-way communication between teachers and students (Black & William, 1998). The SPAQ's ability to distinguish between classes was also established, which was an important contribution of the study. Additionally, scales of attitude to subject and academic efficacy were further validated. High mean scores for scale of attitude to Science describe students positive attitude towards science assessments and is in tune with Klopfer's (1976) second category of structure for evaluating attitudes. Students also demonstrated very high perception of academic efficacy confirming that these students will have high academic motivation (Schunk & Hanson, 1985) persistence (Lyman et al., 1984), memory performance (Berry, 1987), and academic performance (Schunk, 1989).

For gender differences statistically significant differences were found only on one scale of Authenticity at $p < 0.05$ and for all other four scales of the SPAQ and two attitudinal scales no statistically significant differences were recorded. These findings are in conflict with earlier research claims that boys outperformed girls in science (especially physical science) (Casad et al., 2018; Bellar & Gafni, 1996; Kahle & Meece, 1994; Murphy, 1996). This could be place specific where in equal opportunities were being provided to all students in the classroom irrespective of their gender (Fraser et al., 1992; Harding, 1996; Warrington & Younger, 1996). As opposed to results of gender differences for all the scales of the questionnaire statistically significant differences were reported for year level differences, with higher mean scores for Yr 8's and lowest for Yr 10's. The trends of year level differences synchronise with the findings from similar studies (Kim et al., 2000; Koul & Fisher, 2006).

It was found that student perceptual data can be used to identify exemplary teacher and SPAQ was a valid instrument to use for this purpose. The exemplary teachers were identified as those who scored more than one standard deviation above the mean for at least three of the five scales of SPAQ. This resonates with the constructivist view of learning wherein target assertions are clear-cut, students are provided with focused feedback and they are also involved in self and peer assessments (Maulana et al., 2021; Muijs & Reynolds, 2017, p. 1; Sadler, 1989).

Qualitative data added a new rich layer of understanding to already existing knowledge gained through quantitative data. While developing the SPAQ different dimensions of assessment were identified namely, Congruence with planned learning, Authenticity, Student consultation, Transparency and Diversity were identified.

Observations and interview data identified the same dimensions existing within different sections of assessment process. The identified sections namely, learning, curriculum, classroom and assessment, teacher, and student are integral part of assessments. The identified exemplary teachers were engaging constructivist ways of teaching underpinning formulations of formative assessment (Sadler, 1989). The qualitative data identified the importance and role of involving students in assessment task leading to their learning.

Assessment for learning has emerged as central theme in this study. Identified exemplary teachers were found to be very thorough in their teaching, giving students enough time to prepare for an assessment, allowing students freedom to choose from a variety of assessments and were flexible in teaching and assessment. They also demonstrated an in-depth understanding of science topics they were teaching.

This study demonstrates that scales of learning environment can be used in complex studies where many interrelated variables are assessed. By identifying good science teachers and describing what they do in their classrooms, we have an opportunity to use this information in professional development of other interested teachers. This is one of the ways to bring about desired changes in the educational system.

Appendix: Students' Perceptions of Assessment Questionnaire (SPAQ)

Questions in science tests what I know.

My science assignments/tests examines what I do in class.

My assignments/tests are about what I have done in class.

How I am assessed is like what I do in class.

How I am assessed is similar to what I do in class.

I am assessed on what the teacher has taught me.

I am asked to apply my learning to real life situations.

My science assessment tasks are useful in everyday things.

I find science assessment tasks are relevant to what I do outside of school.

Assessment in science tests my ability to apply what I know to real-life problems.

Assessment in science examines my ability to answer every day questions

I can show others that my learning has helped me do things.

In science I am asked about the types of assessment that are used.

I am aware how my assessment will be marked.

I can select how I will be assessed in science.

I have helped the class develop rules for assessment in science.

My teacher has explained to me how each type of assessment is to be used.

I have a say in how I will be assessed in science.

I understand what is needed in all science assessment tasks.

I know what is needed to successfully complete a science assessment task.

I am told in advance when I am being assessed.
 I am told in advance on what I am being assessed.
 I am clear about what my teacher wants in my assessment tasks.
 I know how a particular assessment task will be marked.
 I have as much chance as any other student at completing assessment tasks
 I complete assessment tasks at my own speed.
 I am given a choice of assessment tasks.
 I am given assessment tasks that suit my ability.
 When I am confused about an assessment task, I am given another way to answer it.
 When there are different ways I can complete the assessment.

Scale Allocations:

Congruence with Planned Learning: 1–6
 Authenticity: 7–12
 Student Consultation: 13–18
 Transparency: 19–24
 Diversity: 25–30

References

- Aldridge, J. M., Fraser, B. J., & Huang, T. (1999). Investigating classroom environments in Taiwan and Australia with multiple research methods. *Journal of Educational Research, 93*, 48–57.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.
- Bellar, M., & Gafni, N. (1996). The 1991 international assessment of educational progress in mathematics and sciences: The gender differences perspective. *Journal of Educational Psychology, 88*, 365–377.
- Berry, J. M. (1987). A self efficacy model of memory performance. Paper presented at the annual meeting of the American Educational Research Association, New York.
- Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7–74.
- Casad, B. J., Petzel, Z. W., & Ingalls, E. A. (2018). A model of threatening academic environments predicts women in STEM majors' self-esteem and engagement in STEM. *Sex Roles, 36*(1), 1–20.
- Commonwealth of Australia. (2019). *Advancing women in STEM strategy*.
- Cowie, B. (2005). Student commentary on classroom assessment in science: A sociocultural interpretation. *International Journal of Science Education, 27*(2), 199–214.
- Cronbach, D. J. (1951). Coefficient alpha and internal structure of tests. *Psychometrika, 16*(3), 297–334.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*, 438–481.
- DeVellis, R. F. (1991). *Scale development: Theory and applications* (Applied Social Research Methods Series, Vol. 26). Sage Publications.
- Dorman, J. P., Waldrip, B. G., & Fisher, D. L. (2008). *Using student perceptions of assessment questionnaire (SPAQ) to develop an assessment typology for science classes*. <https://www.semanticscholar.org/paper/Usingstudent-perceptions-of-assessment-%28SPAQ%29-to-Dorman-Waldrip/9a22274855df7791a20fd07f47b7d8ef2e3fdbd0>
- Dorr-Bremme, D., & Herman, J. (1986). *Assessing student achievement: A profile of classroom practices*. Center for the study of Evaluation, UCLA Graduate School of Education.

- Dunkin, M. J., & Biddle, B. J. (1974). *The study of teaching*. Holt, Rinehart & Winston.
- Farenga, S. J., & Joyce, B. A. (1997). What children bring to the classroom: Learning science from experience. *School Science and Mathematics*, 97, 248–252.
- Fraser, B. J. (1978). Development of a test of science-related attitudes. *Science Education*, 62, 509–515.
- Fraser, B. J. (1981). *Test of science-related attitudes: Handbook*. Australian Council for Educational Research.
- Fraser, B. (1994). Research on classroom and school climate. In D. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 493–541). Macmillan.
- Fraser, B. J. (1998). Science learning environments: Assessments, effects and determinants. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education* (pp. 527–564). Kulwer.
- Fraser, B. J., & Wallberg, H. J. (Eds.). (1991). *Educational environments: Evaluation, antecedents and consequences*. Pergamon Press.
- Fraser, B. J., Tobin, K., & Kahle, J. B. (1992). Learning science with understanding: In search of holy grail. *Research in Science and Technological Education*, 10, 65–81.
- Fraser, B. J., Fisher, D. L., & McRobbie, C. J. (1996, April). Development, validation and use of personal and classforms of a new classroom environment instrument. Paper presented at the annual meeting of the American Educational Research Association, New York, USA.
- Gardner, P. L. (1975). Attitudes to science. *Studies in Science Education*, 2, 1–41.
- Goodrum, D., Hackling, M., & Rennie, L. (2001). *The status and quality of teaching and learning in Australian schools*. Department of Education, Training and Youth Affairs.
- Greenfield, T. A. (1996). Gender, ethnicity, science achievement and attitudes. *Journal of Research in Science Teaching*, 33, 901–933.
- Harding, J. (1996). Girls' achievement in science and technology—Implications for pedagogy. In P. F. Murphy & C. V. Gipps (Eds.), *Equity in classrooms* (pp. 111–123). The Falmer Press.
- Harlen, W. (1998). Teaching for understanding in pre-secondary science. In B. Fraser & K. Tobin (Eds.), *International handbook of science education* (pp. 183–198). Kulwer.
- Kahle, J. B., & Meece, J. (1994). Research on gender issues in the classroom. In D. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 542–557). Maxwell Macmillan International.
- Kim, H., Fisher, D., & Fraser, B. J. (2000). Classroom environment and teacher interpersonal behaviour in secondary classes in Korea. *Evaluation and Research in Education*, 14, 3–22. York: Macmillan.
- Klopper, L. E. (1971). Evaluation of learning in science. In B. S. Bloom, J. T. Hastings & G. F. Madaus (Eds.), *Handbook of formative and summative evaluation of student learning*. McGraw-Hill Book Company.
- Klopper, L. E. (1976). A structure for the affective domain in relation to science education. *Science Education*, 60, 299–312.
- Koul, R., & Fisher, D. (2006). An investigation into teacher interaction, perceptions of learning environment cultural differences and science achievement in India. In D. Fisher, D. Zandvliet, I. Gaynor, & R. Koul (Eds.), *Sustainable communities and sustainable environments: Envisioning a role for science, mathematics and technology education* (pp. 323–335). Curtin University of Technology.
- Lewin, K. (1936). *Principals of topological psychology*. McGraw.
- Lyman, R. D., Prentice-Dunn, S., Wilson, D. R., & Bonfilio, S. A. (1984). The effect of success or failure on self-efficacy and task persistence of conduct-disordered children. *Psychology in the Schools*, 21, 516–519.
- Maulana, R., André, S., Helms-Lorenz, M., Ko, J., Chun, S., Shahzad, A., Irnidayanti, Y., Lee, O., de Jager, T., Coetzee, T., & Fadhilah, N. (2021). Observed teaching behaviour in secondary education across six countries: Measurement invariance and indication of cross-national variations. *School Effectiveness and School Improvement*, 32(1), 64–95, 32 p.
- McRobbie, C. J., & Fraser, B. J. (1993). Associations between student outcomes and psychosocial mathematics laboratory environments. *Journal of Educational Research*, 87, 78–85.

- Midgley, C., & Urdan, T. (1995). Predictors of middle school students' use of self-handicapping strategies. *Journal of Early Adolescence, 15*, 389–411.
- Muijs, D., & Reynolds, D. (2017). *Effective teaching* (4th ed.). Sage. Retrieved from <https://www.perlego.com/book/1431675/effective-teaching-pdf>
- Murphy, P. (1996). The IEA assessment of science achievement. *Assessment in Education, 3*, 213–232.
- Newby, M., & Fisher, D. L. (1997). An instrument for assessing the learning environment of a computer laboratory. *Journal of Educational Computing Research, 16*, 179–190.
- Plake, B. S. (1993). Teacher assessment literacy: Teachers' competencies in the educational assessment of students. *Mid-Western Educational Researcher, 6*, 21–27.
- Popham, W. J. (1997). Consequential validity: Right concern-wrong concept. *Educational Measurement: Issues and Practice, 16*(2), 9–13.
- Radnor, H. (1996). *Evaluation of key stage3 assessment in 1995 and 1996* (research report). University of Exeter.
- Reynolds, D. S., Doran, R. L., Allers, R. H., & Agruso, S. A. (1995). *Alternative assessment in science: A teacher's guide*. University of Buffalo.
- Sadler, D. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119–144.
- Schaffner, M., Burry-Stock, J. A., Cho, G., Boney, T., & Hamilton, G. (2000, April). What do kids think when their teachers grade? Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Schunk, D. H. (1989). Self-efficacy and cognitive skill learning. In C. Ames & R. Ames (Eds.), *Research on motivation in education* (Vol. 3, pp. 13–44). Academic.
- Schunk, D. H. (1995). Self-efficacy and education and instruction. In J. E. Maddux (Ed.), *Self-efficacy, adaptation, and adjustment: Theory, research, and application* (pp. 281–303). Plenum.
- Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficacy and achievement. *Journal of Educational Psychology, 77*, 312–322.
- Stiggins, R. (1994). *Student-centered classroom assessment*. Macmillan College Publishing.
- Stolarchuk, E., & Fisher, D. (1999). The effect of using laptop computers on attitude and achievement in science. *Computers in New Zealand Schools, 11*, 40–43.
- Taylor, P. C., Fraser, B. J., & Fisher, D. L. (1997). Monitoring constructivist classroom learning environments. *International Journal of Educational Research, 27*, 293–302.
- Tobin, K. (1998). Issues and trends in the teaching of science. In B. J. Fraser & K. Tobin (Eds.), *International handbook of science education* (pp. 129–151). Kluwer.
- Walberg, H. J. (1976). Psychology of learning environments: Behavioral, structural, or perceptual? *Review of Research in Education, 4*, 142–178.
- Waldrip, B. G., Fisher, D., & Dorman, J. P. (2009). Identifying exemplary science teachers through their student perceptions of assessment process. *Research in Science and Technological Education, 27*(1), 117–129.
- Warrington, M., & Younger, M. (1996). Goals, expectations and motivation: Gender differences in achievement at GSE. *Curriculum, 17*, 80–93.
- Wubbles, T., & Levy, J. (1993). *Do you know what you look like?: Interpersonal relationships in education*. Falmer.
- Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 202–231). Cambridge University Press.

Rekha B. Koul is Associate Professor at STEM Research Group, School of Education, Curtin University, Australia. She has over three decades of teaching and research experience. Her expertise lies in the development, refinement and validation of questionnaires; investigations of the effects of classroom environments on student outcomes; evaluation of educational programs; teacher-action research aimed at improving their environments and evaluation of curriculum and publishes in this area.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 16

Teacher Effectiveness in Multiple Lenses: Secondary Analysis Lessons in the Measures of Effective Teaching Project



James Ko , Zhijun Chen, Jieyan Celia Lei, and Ridwan Maulana 

Abstract Improving teaching quality to enhance learning has become critical for academics, practitioners, and policymakers. However, very few studies compared the same lessons with various classroom observation instruments to examine whether classroom characteristics in different instruments are similar.

This project aimed to conduct a secondary lesson observation analysis on 423 lesson videos selected from 14,000+ lesson videos previously collected in the *Measures of Effective Teaching* (MET) project. The analysis provided new data on the same lessons previously studied by the MET researchers but observed with two instruments. One internationally validated instrument was used in the international project by Maulana et al. (Sch Eff Sch Improv, 1–32, 2021) to explore the generic teaching characteristics in different countries, while the other instrument was developed purposively to characterise the differences between effective and inspiring teaching.

J. Ko (✉)

Department of Education Policy and Leadership, The Education University of Hong Kong,
Tai Po, Hong Kong
e-mail: jamesko@eduhk.hk

Z. Chen

Department of Education, University of Bath, Bath, UK

J. C. Lei

Department of Education, Shaoyang University, Shaoyang, China

Department of Education Policy and Leadership, Education University of Hong Kong,
Tai Po, Hong Kong

R. Maulana

Department of Teacher Education, University of Groningen, Groningen, The Netherlands
e-mail: r.maulana@rug.nl

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_16

339

The results allowed data comparisons across different projects that are not readily comparable because they used various classroom observation instruments. The results informed the relationships between effective and inspiring teaching.

Keywords Teacher effectiveness · Teaching quality · Classroom instrument · Video lesson analysis

1 Introduction

Beliefs about what constitutes ‘good’ or ‘high’ quality practice in teaching can vary markedly for different age groups of students, at other times and in different contexts. ‘Effectiveness’ is a contested term that can evoke strong emotions because perceived effectiveness links with notions of professional competency and high-stakes accountability in some countries. Researchers may question individual teachers’ beliefs about their professional autonomy. Notions of what constitutes high quality or good teaching, or the idea that teaching is an art or a craft rather than a science, are sometimes used to raise concerns with narrower concepts of effectiveness.

Researchers recognise the importance of effective teaching behaviour for student outcomes, but most teachers still struggle to implement complex teaching skills in their daily classroom practices. The Measures of Effective Teaching (MET) project (Bill & Melinda Gates Foundation, 2018) has provided the research communities with the most extensive dataset on classroom observation with an easily accessible video library for secondary data analysis. To our knowledge, this is the first study to connect two large-scale classroom observation studies through secondary data analysis of selected lesson videos with the same instruments.

2 Theoretical Framework

2.1 *Examining Teacher Effectiveness Through Classroom Observations*

Teacher effectiveness research is a branch of educational effectiveness research, focusing mainly on variations in teaching quality on student outcomes. Value-added measures, classroom observations, and student surveys are familiar sources of information and data about teachers’ behaviour and classroom practices that can be drawn upon to provide evidence to inform our understanding of teacher effectiveness (Bacher-Hicks et al., 2019). If student outcomes are the essential criteria for teacher effectiveness, the question remains about what kinds of outcomes,

objectives, and goals can be achieved by teachers and schools. We clearly cannot go on endlessly adding more objectives and more content for teachers and schools and still expect them to succeed. Teacher effectiveness beyond the classroom level (e.g., Cheng, 1996; Cheng & Tsui, 1998) is not practically appealing to practitioners because such a conceptualisation could obscure the focus of the teacher's role and duties in teaching. In practice, teachers and schools often prefer to restrict teacher evaluation to specific objectives in teaching.

For evaluating teaching quality, while gains in cognitive and non-cognitive domains of student achievement are tentative, a plea for meeting cognitive purposes and obtaining higher academic attainments as the criteria for the effectiveness of education in schools often sounds more appealing. Bacher-Hicks et al. (2019) argued for value-added measures as unbiased predictors of teacher performance in experimental conditions where students were assigned randomly to different classrooms. However, value-added measures are not unbiased as assumed because they tend to shift when different tests assess student achievements (Grossman et al., 2014).

While a classroom observation approach cannot adjust for classroom composition, it has two obvious advantages apart from easily accessible applications in naturally occurring settings. First, it allows ready comparisons across grades and subjects without relying on reliable, standardised tests. Second, a classroom observation approach looks at teacher effectiveness from a different angle by allowing the observers or evaluators to associate the observed behaviours with various aspects of the student learning process, such as student engagement in class and students' self-reported behaviours or learning characteristics (e.g., Clunies-Ross et al., 2008; Helmke et al., 1986; Virtanen et al., 2015).

2.2 *Comparisons of Classroom Observation Instruments*

Classroom observation is a powerful method to collect data on teacher behaviours in class. Numerous sources of information and data about teachers' behaviour and classroom practices can be drawn upon to provide evidence to inform our understanding of teacher effectiveness. A standard method in a classroom observation approach to teacher effectiveness research is to observe different teachers' teaching practices by independent observers. For example, this paper compared the results obtained from different high-inference instruments to capture aspects of teaching dimensions hypothesised to be operated at the classroom level.

We can compare different classroom observation instruments of similar nature (i.e., for generic teaching behaviours) for different lessons in a single study (Day et al., 2008; Kington et al., 2014), different classroom observation instruments of similar nature for the same lessons in a single study (e.g., Ko, 2010; Ko et al., 2015; Lei et al., 2023; Sammons et al., 2014, 2016), and different classroom observation instruments of different nature (i.e., effective vs inspiring teaching behaviours) for different lessons in a single study (Ko et al., 2019a, b, 2016; Zhao & Ko, 2022).

However, the measurement strategy of teacher effectiveness in the MET project was unique as it compared different classroom instruments that differed in specificity. It involved comparisons of generic teacher behaviours by the *Framework for Teaching* (Danielson, 2013) and the *Classroom Assessment Scoring System* (CLASS) (Pianta et al., 2012) and of subject-specific ones by the *Mathematical Quality of Instruction* (MQI) (Hill et al., 2008; The Learning Mathematics for Teaching Project, 2011), the *Protocol for Language Arts Teaching Observations* (PLATO), the *Quality of Science Teaching* (QST) (Schultz & Pecheone, 2014), and the *UTeach Teacher Observation Protocol* (UTOP) (Walkington & Marder, 2014, 2018).

The challenges of developing and comparing quality teacher observation systems lie in establishing rater reliability and making the instruments more generalisable across contexts (Hill et al., 2012; Liu et al., 2019). For example, despite its wide application, the CLASS was not adequately validated without revisions in Hafen et al. (2015), where the lesson videos were collected in various projects. Wallace et al. (2020) also reported that the CLASS failed to discriminate classroom management quality, with most teachers' scores clustering around the most positive ranges of effectiveness. The present study differed from the heuristic comparison of classroom observation instruments by Bell et al. (2019) in that we observed the same lessons with different instruments. Bell et al.'s (2019) comparison was crude and non-quantitative, as all instruments they compared shared ten similar teaching dimensions.

Secondary data analysis on the MET data should provide quantitative evidence for instrument comparisons, but to date, we still cannot find any study exploring this. We intended to fill this gap with this study and were motivated to conduct secondary data analysis with the same classroom observation instrument in the international collaborative ICALT3 project (Maulana et al., 2021) so that the new data on the selected sample would form a part of the enlarged study to inform the measurement invariance of teaching quality (Krammer et al., 2020; Maulana et al., 2021).

2.3 Video Lesson Analysis¹

The TIMSS 1995 video study by Stigler et al. (1999) was a pioneer and exemplar in using video data to explore teaching characteristics and patterns cross-culture beyond qualitative coding to provide quantitative analysis for hypothesis testing

¹We deliberately left out discussing the OECD TALIS video study (McCann et al., 2020; OECD, 2017) for this section because the study was still on-going by the time of writing. Although a very elaborated observation instrument was developed for the study, it has not been applied to any other study beyond the OECD. We also excluded TEACH (World Bank, 2019) for its limited application in research to date, but a chapter by (Lei et al., 2023, this volume) that compares TEACH and ICALT can be found in this volume.

(Jacob et al., 1999; Stigler et al., 2000). The initial sample included 231 mathematics lessons from Germany, Japan, and the United States, selected from a nationally representative sample of eighth-grade students and classrooms participating in the 1994–95 TIMSS assessments. The TIMSS 1999 video study expanded to include Australia, the Czech Republic, Hong Kong SAR, the Netherlands, Switzerland, and the United States (Hiebert et al., 2003). In the video study on science teaching, the participating countries included Australia, the Czech Republic, Japan, the Netherlands, and the United States (Roth et al., 2006). However, no observation instruments were developed or adopted for observations in these studies. Only a portion of the lesson videos are publicly available for secondary data analysis, so our purpose should provide new data for the ICALT3 with the MET data.

Apart from the MET project, only a few studies in the literature used lesson videos to conduct lesson observation to inform teaching practice and performance (e.g., Hafen et al., 2015; Ko et al., 2015, 2016). Secondary data analysis makes instrument comparisons feasible if a lesson is videotaped for observation, as in the MET project, providing opportunities to observe the same lessons again at different times and research contexts.

3 Methods

3.1 Data Collection

The current study used both the original data of the CLASS and new observational data using two classroom observations to compare classroom characteristics.

3.2 Raters

The second and third authors conducted the majority of the lessons. The third author assisted the second author as a research assistant to use ICALT in another project (Lei et al., 2023). When these research assistants shared the secondary video analysis, they passed the training session and conducted two calibrations. One English for Language Arts (ELA) lesson and one Math lesson were observed and scored twice by each rater in each calibration. After each calibration, they conducted an interrater reliability test and proceeded with the lesson observation when Krippendorff's alpha (2004) increased from .52 to .73 for the ELA lesson and from .55 to .82 for the ELA Math lesson.

3.3 Video Samples

3.3.1 Original Lesson Videos of the MET Study

The *Measurement of Effective Teaching* (MET) project was a large project funded by the Bill & Melinda Gates Foundation (2018). Around 2700 teachers from 10 districts in the United States teaching science, English, and math across 4–9 grades participated from 2009–2010 and 2010–2011 (Bill & Melinda Gates Foundation, 2018). Each teacher was videotaped during the lessons one to four times over a year. After training, the lessons were divided into segments and coded in 20-min segments by their administrator and peer observers using different classroom observation instruments. Despite its scale, teachers, classrooms, schools and districts in the MET project were not randomised.

4 Current Secondary Data Analysis

Among these different instruments, the *CLASS* was used for all lessons in the MET project and the most studied instrument outside the U.S.A. (e.g., Taut et al., 2019 in Chile; Pöysä et al., 2019 in Finland; Havik & Westergård, 2020 in Norway). The *CLASS* was assumed to be a reliable reference for selecting lesson videos for secondary data analysis with two new classroom observation instruments, the *International Comparative Analysis of Learning and Teaching* (ICALT) and the *Comparative Analysis of Effective Teaching and Inspiring Teaching* (CETIT). Thus, in this study, we selected four hundred twenty-three lessons proportional to the staine distribution (Clark-Carter, 2005) of the percentiles of the aggregated mean scores of the various teaching dimensions of the *CLASS*.² We also limited the sample to secondary school lessons (i.e., 7–9) and English and mathematics only. Two lessons were excluded due to low video quality. Three trained raters observed nine lessons for calibrations first and started secondary observations after inter-rater reliability was over 90%. Each observer was assigned randomly to observe different lessons. The total numbers of segment, video, rater, and teacher are summarised in Table 16.1.

²The distribution of *CLASS* scores in the MET project was normal. Thus, using staines to select sample lessons for the secondary analysis retained a similar normal distribution.

Table 16.1 The total numbers of segment, video, rater and teacher of the original data in the MET project and 423 chosen in this project

	Original MET	Current secondary data analysis
Segment	3	1–3
Lesson	14,179	423
Rater	419	3
Teacher	1594 (4 lessons)	217 (1 lesson); 103 (2 lessons)
Year	1st (2010); 2nd (6294)	1st (318); 2nd (105)
Grade	4th–9th	7th–9th
Subject	2	2

4.1 Instruments

4.1.1 CLASS

The *CLASS* in the MET project has an additional dimension, *Instruction Dialogue*, in addition to its original version with ten³ dimensions of teaching quality: *Positive Climate*, *Negative Climate*, *Teacher Sensitivity*, *Adolescent Perspectives*, *Behaviour Management*, *Productivity*, *Instructional Learning Formats*, *Content Understanding*, *Analysis and Inquiry*, *Quality of Feedback*, and one dimension of *Learner Engagement* (Pianta & Hamre, 2009). Each lesson was divided into one, two, or three segments, each rated independently by a different rater on a 7-point Likert scale representing low to high levels.

4.1.2 ICALT

Originally developed as an instrument for inspection to capture generic teaching behaviours (van de Grift, 2007, 2014), the *ICALT* has expanded into thirty-two high-inference teaching indicators categorised into six domains: *Safe and stimulating learning climate*, *Efficient organisation*, *Clear and structured instructions*, *Intensive and activating teaching*, *Adjusting instructions and learner processing to inter-learner differences*, and *Teaching learning strategies*. The *ICALT* also contained a three-item (e.g., ‘...take an active approach to learn’) student learning domain to document learner engagement during classroom observations. Three observers completed classroom observation for each lesson and rated the items based on teachers’ performance on a 4-point scale, from ‘mostly weak’ to ‘mostly strong.’

³In Pianta and Hamre (2009), there was a dimension of *Procedures and skills*, which was not in the MET.

4.1.3 CETIT

Based on the teaching aspects characterised as inspiring teaching by Sammons et al. (2014), Ko et al. (2016) used the Delphi method to finalise and validate the *CETIT*. This new high-inference classroom observation instrument consisted of sixty-eight descriptive statements that included effective and inspiring teaching domains. According to Ko et al. (2016), inspiring teaching includes four aspects: *Flexibility*, *Teaching reflective thinking*, *Innovative teaching*, and *Teaching collaborative learning*. Teaching behaviours corresponding to these inspiring teaching domains include “*The teacher allowed options for students in their seatwork*,” “... *asked students to comment on his/her viewpoint*,” “... *used ICT in teaching*,” and “... *told students how to share their work in a task*.” While *Teaching reflective thinking* and *Teaching collaborative learning* were two distinctive classroom practices in the *CETIT*, they were conceived as a single characteristic by Sammons et al. (2014). *Dimensions Assessment for learning* and *Professional Knowledge and expectations* are two unique teaching aspects in the *CETIT* (i.e., not found in the *CLASS* or the *ICALT*). They were found to cluster with other teaching domains of effective teaching (Ko et al., 2016, 2019a, b). For this study, two new dimensions, *Engagement in exploratory learning* and *Engagement in knowledge consolidation*, developed by Piburn and Sawada (2000), were adopted to test whether the learner dimensions in different instruments might favour the teaching dimensions of the classroom observation instruments to which they belong.

4.2 Data Analysis

For all three instruments, the means, standard deviations, and reliability tests were conducted in SPSS 20. Confirmatory factor analyses (CFA) were conducted in MPlus 7. The original three-factor model of the *CLASS* was tested first, followed by one-factor and two-factor models for comparison. For the *ICALT*, a six-factor model was tested with the theoretical structure. Three CFA models were tested on the *CETIT*: (a) an eight-factor model on effective teaching, (b) a four-factor model on inspiring teaching, and (c) a 12-factor full model. Multiple good fit indices were selected as the criteria suggested by Tabachnick et al. (2007) for evaluating the CFA models: (a) the Root mean square error of approximation (RMSEA) below .08, (b) a Comparative fit index (CFI) above .95, (c) standardised root mean square residual under .08, and (d) χ^2/df to be under 2.

5 Findings

5.1 Descriptive Statistics

5.1.1 CLASS

The overall results shown in Table 16.2 are consistent with the CLASS results in the literature. *Instructional support* was the weakest domain. At the dimension level, the average scores were relatively low for *Negative Climate* ($M = 1.47$, $SD = .63$) and *Analysis and Inquiry* ($M = 2.42$, $SD = .90$). In contrast, Dimensions *Behavior Management* ($M = 5.72$, $SD = .98$) and *Productivity* ($M = 5.54$, $SD = .93$) were scored relatively higher than all other dimensions. Table 16.2 indicated that the reliability for each domain, *Emotional Support*, *Classroom Organisation*, or *Instructional Support*, was acceptable as a subscale and the full-scale CLASS ($\alpha > .7$). There were no reliability scores for dimensions because they were single indicators.

5.1.2 ICALT

For ICALT, the means of *Adjusting Instructions and Learner Processing to Inter-Learner Differences* ($M = 1.68$, $SD = .42$) and *Teaching Learning Strategies* ($M = 1.45$, $SD = .40$) were low because they were rare in the sampled lessons. The reliability test results indicated a high level of internal consistency for the full scale of ICALT ($\alpha = .87$). Still, as depicted in Table 16.3, half of the ICALT domains have

Table 16.2 Means, standard deviations and Cronbach's Alphas (α) of CLASS

Dimension/Domain	Mean	SD	Cronbach
Positive climate	4.14	.97	NA
Negative climate	1.47	.63	NA
Teacher sensitivity	3.87	.97	NA
Regard for adolescent perspectives	2.93	1.03	NA
Behaviour management	5.72	.98	NA
Productivity	5.54	.93	NA
Instructional learning formats	3.85	.91	NA
Content understanding	3.59	.98	NA
Analysis and inquiry	2.42	.90	NA
Quality of feedback	3.21	1.02	NA
Instructional dialogue	2.94	1.06	NA
Student engagement	4.50	.92	NA
<i>Emotional support</i>	4.12	.70	.76
<i>Classroom organization</i>	5.04	.75	.72
<i>Instructional support</i>	3.04	.87	.90
Full scale	4.02	.68	.90

Table 16.3 Mean, standard deviation and Cronbach's Alpha (α) of ICALT

Dimension	Mean	SD	Cronbach
Safe and stimulating learning climate	3.04	.45	<u>.58</u>
Efficient organisation	3.30	.40	<u>.65</u>
Clear and structured instructions	2.88	.41	.76
Intensive and activating teaching	2.29	.41	<u>.56</u>
Adjusting instructions and learner processing to inter-learner differences	1.68	.42	<u>.49</u>
Teaching learning strategies	1.45	.40	.70
Learner engagement	2.93	.49	.79
Full scale	2.40	.30	.87

Table 16.4 Mean, standard deviation and Cronbach's Alpha (α) of CETIT

Dimension	Mean	SD	Cronbach
Enthusiasm for teaching	3.17	.49	.84
Positive relationships with students	2.72	.47	.72
Purposeful and relevant teaching	1.89	.42	<u>.57</u>
Safe classroom climate	2.86	.46	<u>.58</u>
Stimulating learning environment	1.43	.39	<u>.23</u>
Positive classroom management	3.37	.45	.73
Assessment for learning	2.18	.35	<u>.61</u>
Professional knowledge and expectations	2.88	.40	<u>.60</u>
Flexibility	1.13	.27	<u>.58</u>
Teaching reflective thinking	1.35	.34	.70
Teaching collaborative learning	1.43	.54	<u>.68</u>
Innovative teaching	1.32	.35	<u>.59</u>
Engagement in exploratory learning	1.68	.31	<u>.54</u>
Engagement in knowledge consolidation	2.33	.42	<u>.49</u>
Full scale	2.10	.26	.93

reliability below .7, the threshold acceptable in education research (Taber, 2018): *Safe and Stimulating Learning Climate* ($\alpha = .58$), *Intensive and Activating Teaching* ($\alpha = .56$) and *Adjusting Instructions and Learner Processing to Inter-Learner Differences* ($\alpha = .49$).

5.1.3 CETIT

The result suggested that the full scale with all 68 items was highly consistent ($\alpha = .93$). The result also indicated good reliabilities in most of the CETIT dimensions. Besides, there was an unacceptable internal consistency of the subscale *Stimulating Learning Environment* ($\alpha = .23$), with a relatively lower score average ($M = 1.43$, $SD = .39$). The four subscales with reliability close to the .6-threshold included *Flexibility* ($\alpha = .58$), *Purposeful and Relevant Teaching* ($\alpha = .57$), *Safe Classroom Climate* ($\alpha = .58$), and *Innovative Teaching* ($\alpha = .59$) (Table 16.4).

6 Correlations of Factors of Three Instruments

Table 16.5 displays the Pearson correlation coefficients of the teacher dimensions, learner engagement dimensions, and the whole scale. As correlations are sensitive to sample size, we should focus on the association's magnitude or strength. In general, a coefficient between .4 and .6 indicates a moderate strength. While a value above .6 suggests a strong association, a value between .2 and .4 is weak to mild. Values below .2 are considered weak even though the correlation may be statistically significant.

Most teaching dimensions of the *CLASS* were correlated significantly only with other dimensions of the same scale, but teaching dimensions of *ICALT* and *CETIT* correlated with other dimensions of each other scale. All eleven *CLASS* dimensions suggested weak or no correlations with the *ICALT* and *CETIT* dimensions. In the *ICALT*, the result indicated that the domain *Teaching Learning Strategies* did not correlate with three domains in the *ICALT*: *Safe and Stimulating Learning Climate*, *Efficient Organisation*, and *Clear and Structured Instructions*.

In the *CETIT*, the dimension *Innovative Teaching* showed no correlation with the other nine dimensions, except for *Flexibility* ($r = .271, p < .01$) and *Teaching Reflective Thinking* ($r = .280, p < .01$). All three domains were classified as inspiring teaching practices. In contrast, other *CETIT* dimensions were correlated significantly with most *ICALT* dimensions. Comparing the subscales of student engagement in the *CLASS*, *ICALT*, and *CETIT*, *Learner Engagement* in the *ICALT* showed stronger correlations with more teaching dimensions, six in the *CLASS*, five in the *ICALT*, and ten in the *CETIT*. *Learner Engagement* in the *ICALT* was also weakly associated with *Student Engagement* in the *CLASS* ($r = .239, p < .01$), *Engagement in Exploratory Learning* ($r = .481, p < .01$) and *Engagement in Knowledge Consolidation* ($r = .608, p < .01$) in the *CETIT*.

6.1 Comparing Confirmatory Factor Models of Three Instruments

6.1.1 CLASS

Except for the original three-factor models of the *CLASS*, the one-factor and two-factor models were also built up to investigate a better factor structure of the *CLASS* based on the sampled lessons. In all three models, the two-factor model showed a relatively better model fit than the other one-factor model and the original three-factor model. The one-factor model of the *CLASS* suggested poor model fit to the data, $\chi^2(54) = 846.491, p < .001, CFI = .768, RMSEA = .186$, but interestingly, the theoretical three-factor model of the *CLASS* had the worst fit indices, $\chi^2(41) = 774.629, p < .001, CFI = .761, RMSEA = .205$. In contrast, the two-factor

Table 16.5 Correlations between CLASS, ICALT & CETIT

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34				
1 CLASS_Pedagogic Climate	1																																					
2 CLASS_Supportive Climate	0.29**	1																																				
3 CLASS_Instructional Quality	0.15**	0.58**	1																																			
4 CLASS_Role for Adults/Prerequisites	0.07*	0.10**	0.09**	1																																		
5 CLASS_Behavioral Management	0.28**	0.68**	0.20**	0.11**	1																																	
6 CLASS_Pedagogical	0.14**	-0.12**	-0.32**	0.16**	0.46**	1																																
7 CLASS_Instructional Learning Formats	0.59**	0.15**	0.61**	0.69**	0.30**	0.39**	1																															
8 CLASS_Content Understanding	0.65**	0.17**	0.54**	0.57**	0.48**	0.72**	0.1	1																														
9 CLASS_Analysis and Inquiry	0.51**	-0.11*	0.40**	0.29**	0.19**	0.25**	0.55**	0.47**	1																													
10 CLASS_Quality of Feedback	0.51**	-0.12**	0.69**	0.59**	0.20**	0.21**	0.65**	0.79**	0.63**	1																												
11 CLASS_Instructional Dialogue	0.57**	-0.16**	0.59**	0.70**	0.21**	0.29**	0.43**	0.69**	0.79**	0.56**	1																											
12 ICALT_0-6th and remaining learning climate	0.20**	-0.02	0.27**	0.20**	0.04*	0.06	0.22**	0.17**	0.25**	0.21**	0.1	1																										
13 ICALT_0-6th and remaining instruction	0.13**	-0.02	0.17**	0.13**	0.14**	0.20**	0.15**	0.13**	0.24**	0.1	0.24**	0.1	1																									
14 ICALT_7-9th and remaining learning climate	0.19**	-0.04	0.24**	0.17**	0.19**	0.19**	0.20**	0.22**	0.22**	0.22**	0.19**	0.20**	0.19**	1																								
15 ICALT_7-9th and remaining instruction	0.19**	-0.04	0.24**	0.17**	0.19**	0.19**	0.20**	0.22**	0.22**	0.22**	0.19**	0.20**	0.19**	0.1	1																							
16 ICALT_10-12th and remaining learning climate	0.15**	0.03*	0.24**	0.06*	0.02*	0.06	0.03	0.06	0.12*	0.14**	0.02*	0.17**	0.22**	0.24**	0.62**	1																						
17 ICALT_10-12th and remaining instruction	0.15**	0.03*	0.24**	0.06*	0.02*	0.06	0.03	0.06	0.12*	0.14**	0.02*	0.17**	0.22**	0.24**	0.62**	0.62**	1																					
18 ICALT_10-12th and remaining learning climate	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
19 ICALT_10-12th and remaining instruction	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
20 ICALT_10-12th and remaining learning climate	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
21 ICALT_10-12th and remaining instruction	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
22 ICALT_10-12th and remaining learning climate	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
23 ICALT_10-12th and remaining instruction	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
24 ICALT_10-12th and remaining learning climate	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
25 ICALT_10-12th and remaining instruction	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
26 ICALT_10-12th and remaining learning climate	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
27 ICALT_10-12th and remaining instruction	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
28 ICALT_10-12th and remaining learning climate	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
29 ICALT_10-12th and remaining instruction	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
30 ICALT_10-12th and remaining learning climate	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
31 ICALT_10-12th and remaining instruction	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
32 ICALT_10-12th and remaining learning climate	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
33 ICALT_10-12th and remaining instruction	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
34 ICALT_10-12th and remaining learning climate	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		
35 ICALT_10-12th and remaining instruction	0.03*	0.07*	0.14**	0.01*	0.07*	0.12**	0.09*	0.17**	0.12**	0.14**	0.12**	0.03	0.06	0.08	0.14**	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08		

* Correlation is significant at the .05 level

** Correlation is significant at the .01 level

model of the CLASS had the best but still unacceptable fit indices, $\chi^2(43) = 358.418$, $p < .001$, CFI = .897, RMSEA = .131.

6.1.2 ICALT and CETIT

Relatively speaking, the results indicated a poor model fit for the six-factor model of the *ICALT*, $\chi^2(449) = 2823.249$, $p < .001$, CFI = .558, RMSEA = .112, while all three CFA models of the *CETIT* suggested relatively better model fits than the *ICALT* ones. The eight-factor model of effective practices in the *CETIT*, $\chi^2(1091) = 4796.771$, $p < .001$, CFI = .618, RMSEA = .09, have better fit indices than those of the four-factor model of inspiring practices, $\chi^2(149) = 782.373$, $p < .001$, CFI = .659, RMSEA = .1, except for the SRMR. However, the full 12-factor model has overall the best fit (except for CFI) among all CFA models with $\chi^2(2144) = 802.596$, $p < .001$, CFI = .572, and RMSEA = .08.

7 Discussions

7.1 Teaching Effectiveness in Different Lens

This secondary analysis was intended to examine teacher effectiveness by comparing different classroom observation instruments. Theoretically, CLASS and ICALT have similar teaching dimensions, but our results showed that ICALT and CETIT were more closely correlated. We could not rule out that this closer relationship was a halo effect of the rater effect because the same raters rated them. While all three scales were reliable, some of the individual dimensions of ICALT and CETIT were internally inconsistent, contrary to the latest research (e.g., Ko & Li, 2020; Maulana et al., 2021). The most puzzling findings were the insignificant relationship of the factors in the confirmatory factor analyses of the three instruments in Table 16.6.

Table 16.6 Model fit indices of confirmatory factor models of CLASS, ICALT and CETIT

	CLASS			ICALT	CETIT		
	1-factor	2-factor	3-factor	6-factor	Effective-8 factor	Inspire-4 factor	12-factor
TLI	.716	.868	.679	.512	.593	.620	.546
CFI	.768	.897	.761	.558	.625	.659	.572
RMSEA	.186	.131	.205	.112	.095	.1	.08
SRMR	.105	.068	.154	.134	.113	.099	.112
χ^2/df	15.675	8.335	18.893	6.287	4.829	5.219	3.741

7.2 *Validity and Reliability of Instruments*

The major limitation of the current study was the poor validity and reliability of the instruments. Though the CLASS and ICALT have been validated in many international contexts, we failed to validate them in the selected sample. We do not intend to provide arguments for retaining the models with poor fit indices nor discuss strategies to modify the model to obtain an acceptable fit because this would go beyond the purpose of this paper. To our surprise, the two-factor model showed a better fit than the theoretical three-factor model. However, similar results were reported by Hafen et al. (2015), who found their bi-factor model fitted the MET data better than the original three-factor model. It is beyond this book chapter's scope to explore a possible revised three-factor model. Still, the results suggested that the CLASS could be inherently unstable because the *Instruction Support* domain is empirically more distinctive than the other domains.

Regarding instrument comparison, the CFA results favoured the CETIT slightly, more for its effective teaching component than its inspiring teaching component. Further studies on the relative significance of individual teaching dimensions (or subscales) will help us further teacher effectiveness research from scale or instrument development to teacher development conceptualising teaching practices ranked by difficulties (Ko et al., 2016).

We are also surprised that the reliability scores of some of the subscales of the ICALT and CETIT were unacceptably low. These results differed much from what we found in our previous projects (Ko et al., 2016, 2019a, b; Maulana et al., 2021). These results might raise concerns over the reliability of the raters' judgements. Given the high-inference nature of classroom observation instruments, ratings are expected to be evaluative. Though we had trained our raters and did calibration to minimise subjective biases in our observations, halo effects might affect the raters' judgements, making the results of the ICALT and CETIT more similar to each other than the CLASS. However, we are more inclined to suspect that this might be a side effect of a biased sample (see below). Still, further analyses to explore any rater effects seem wanting.

7.3 *Limitations with the Original MET Sample*

Conducting classroom observation or teacher evaluation research has been challenging because teacher evaluation is always a sensitive matter for practitioners. The MET lessons were not naturalistic and subject to self-selection bias because teachers and schools provided lesson video clips. There was little control over the quality of the recording and the settings. The video quality might affect the raters' judgments of student engagement as students were often off the screen. However, the secondary video data analysis could be a strength because this allowed other researchers to build up a video-based lesson database with other instruments.

Since we suspected there might be a problem using aggregate averages as references to select our lesson sample, we did another CFA with all the MET lessons to establish the scale validity, but the fit indices were also disappointing. We could not find any report concerning CLASS validation in the MET in its documentation or the literature. We could not identify what characteristics in the entire MET sample and our lesson sample might have caused the inadequate validation. Our assumption that the validations of the *ICALT* and *CETIT* were much affected by some unknown biased sample selection may not be justifiable as it seems. Moreover, we have not conducted further analyses to check systematic biases regarding teacher, school and district characteristics, as we assumed they would be marginal compared to variations in teaching quality.

7.3.1 Significance and Implications

Studying teachers' classroom practices and their effects is essential for teacher development and school improvement. We regard this study's significant implication in indicating the relative strengths and areas for teaching improvements (i.e., flexibility, innovative teaching, adjusting instructions and learner processing to inter-learner differences, teaching learning strategies). Future training on the *CETIT* and *ICALT* as reflective tools may benefit practitioners.

Despite the limitations discussed, this study provides data for instrument comparisons. Some teaching practices are comparable across instruments. Instrument comparison was already an essential focus in the MET project, which included six observation instruments, including more generic by nature, the *CLASS*, and more subject-specific ones like the *PLATO*, *MQI*, *QST*, and *UTOP*. Future research should extend comparisons to these subjects-specific instruments.

The secondary data analysis was a cost-effective strategy to connect two independent studies, the MET and *ICALT3* projects. The secondary data analysis could be done because the lessons were videotaped, providing opportunities to observe the same lessons again at different times and in research contexts. However, secondary data analysis is also limited by the quality of the original sample also limits secondary data analysis as the researchers who conduct secondary data analysis can do little to rectify flaws in the data collection processes.

8 Conclusion

It is tricky and controversial to define effective teaching or teaching effectiveness. Effective teaching requires criteria for effectiveness. The criteria implied in the various teaching dimensions in the *CLASS*, *ICALT* and *CETIT* refer to education objectives in general and teaching in particular. Visions about these criteria result from a political and societal debate, but educational professionals, teachers and schools can also participate in classroom observations. Going beyond identifying

effective classroom practice characteristics, we have uncovered the similarities and variations across teaching dimensions in different instruments. It was surprising that the CLASS could not be validated in our sample as in the original MET dataset. Despite limitations in the validity and reliability of the samples, we consider that our attempt to provide data for the ICALT3 project is at least partially fulfilled.

References

- Bell, C. A., Dobbelaer, M. J., Klette, K., & Visscher, A. (2019). Qualities of classroom observation systems. *School Effectiveness and School Improvement*, 30(1), 3–29.
- Bill & Melinda Gates Foundation. (2018). *Measures of effective teaching project: Frequently asked questions*. Retrieved from <http://k12education.gatesfoundation.org/blog/measures-of-effective-teaching-project-faqs/>
- Bacher-Hicks, A., Chin, M. J., Kane, T. J., & Staiger, D. O. (2019). An experimental evaluation of three teacher quality measures: Value-added, classroom observations, and student surveys. *Economics of Education Review*, 73, 101919.
- Boston, M., Bostic, J., Lesseig, K., & Sherman, M. (2015). A comparison of mathematics classroom observation protocols. *Mathematics Teacher Educator*, 3(2), 154–175.
- Cheng, Y. C. (1996). Total teacher effectiveness: new conception and improvement. *International Journal of Educational Management*, 10(6), 7–17.
- Cheng, Y. C., & Tsui, K. T. (1998). Research on total teacher effectiveness: Conception strategies. *International Journal of Educational Management*, 12(1), 39–47.
- Clark-Carter, D. (2005). Stanine Scores. In P. R. Rosenbaum, B. S. Everitt, & D. Howell (Eds.), *Encyclopaedia of statistics in behavioural science* (Vol. 4, p. 1893). Wiley.
- Clunies-Ross, P., Little, E., & Kienhuis, M. (2008). Self-reported and actual use of proactive and reactive classroom management strategies and their relationship with teacher stress and student behaviour. *Educational Psychology*, 28(6), 693–671.
- Danielson, C. (2013). *The framework for teaching evaluation instrument*. Danielson Group.
- Day, C., Sammons, P., Kington, E., & Regan, E. K. (2008). *Effective classroom practice: A mixed-method study of influences and outcomes*. ESRC.
- Grossman, P., Cohen, J., Ronfeldt, M., & Brown, L. (2014). The test matters: The relationship between classroom observation scores and teacher value-added on multiple types of assessment. *Educational Researcher*, 43(6), 293–303.
- Hafen, C. A., Hamre, B. K., Allen, J. P., Bell, C. A., Gitomer, D. H., & Pianta, R. C. (2015). Teaching through interactions in secondary school classrooms: Revisiting the factor structure and practical application of the classroom assessment scoring system–secondary. *The Journal of Early Adolescence*, 35(5–6), 651–680.
- Havik, T., & Westergård, E. (2020). Do teachers matter? Students' perceptions of classroom interactions and student engagement. *Scandinavian Journal of Educational Research*, 64(4), 488–507.
- Helmke, A., Schneider, W., & Weinert, F. E. (1986). Quality of instruction and classroom learning outcomes: The German contribution to the IEA classroom environment study. *Teaching and Teacher Education*, 2(1), 1–18.
- Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., Chui, A. M., Wearne, D., Smith, M., Kersting, N., Manaster, A., Tseng, E., Etterbeek, W., Manaster, C., Gonzales, P., & Stigler, J. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study* (NCES (2003-013)). U.S. Department of Education, National Center for Education Statistics.

- Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, 26(4), 430–511.
- Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalisability study. *Educational Researcher*, 41(2), 56–64.
- Jacobs, J. K., Kawanaka, T., & Stigler, J. W. (1999). Integrating qualitative and quantitative approaches to the analysis of video data on classroom teaching. *International Journal of Educational Research*, 31(8), 717–724.
- Kington, A., Sammons, P., Brown, E., Regan, E., Ko, J., & Buckler, S. (2014). *Effective classroom practice*. Open University Press.
- Ko, J. Y. O. (2010). Consistency and variation in classroom practice: A mixed-method investigation based on case studies of four EFL teachers of a disadvantaged secondary school in Hong Kong (Doctoral dissertation, University of Nottingham).
- Ko, J., & Li, W. (2020, April). *Effective teaching and inspiring teaching in different learning environments: Evidence from cluster analysis and SEM*. The 2020 Annual meeting of American Educational Research Association (AERA), Virtual Conference, USA.
- Ko, J., Chen, W., & Ho, M. (2015). *Teacher effectiveness and goal orientation* (School Report No. 1). Hong Kong Institute of Education.
- Ko, J., Sammons, P., & Kyriakides, L. (2016). *Developing instruments for studying inspiring teachers and their teaching practice in diverse contexts. Final report*. The Education University of Hong Kong.
- Ko, J., Fong, Y., & Xie, A. (2019a, August). *Effective and inspiring teaching in Math and Science classrooms: Evidence from systematic classroom observation and implications on STEM education*. WERA Annual Meeting in Tokyo.
- Ko, J., Sammons, P., Maulana, R., Li, W., & Kyriakides, L. (2019b, April). *Identifying inspiring versus effective teaching: how do they link and differ?*. The 2019 Annual meeting of American Educational Research Association (AERA), Toronto, Canada.
- Krammer, G., Pflanzl, B., Lenske, G., & Mayr, J. (2020). Assessing quality of teaching from different perspectives: Measurement invariance across teachers and classes. *Educational Assessment*, 26, 1–16.
- Learning Mathematics for Teaching Project. (2011). Measuring the mathematical quality of instruction. *Journal of Mathematics Teacher Education*, 14, 25–47.
- Lei, J., Chen, Z., & Ko, J. (2023). Differences in perceived instructional quality of the same classrooms with two different classroom observation instruments in China: Lessons learned from qualitative analysis of four lessons using TEACH and ICALT. In R. Maulana, M. Helms-Lorenz, & R. M. Klassen (Eds.), *Effective teaching around the world*. Springer.
- Liu, S., Bell, C. A., Jones, N. D., & McCaffrey, D. F. (2019). Classroom observation systems in context: A case for the validation of observation systems. *Educational Assessment, Evaluation and Accountability*, 31(1), 61–95.
- Maulana, R., Andre, S., Helms-Lorenz, M., Ko, J., Chun, S., & Shahzad, A. (2021). Observed teaching behaviour in secondary education across six countries: Measurement invariance and indication of cross-national variations. *School Effectiveness and School Improvement*, 32(1), 64–95.
- McCann, E., Riggall, A., Ingram, J., & Lindorff, A. (2020). *TALIS video study: Technical report*. DFE-RR109. Department for Education, UK.
- OECD. (2017). *TALIS 2018 video study and global video library on teaching practices*. OECD.
- Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, 38(2), 109–119.
- Pianta, R. C., Hamre, B. K., & Mintz, S. (2012). *Classroom Assessment Scoring System (CLASS) Manual, Secondary*. Teachstone.
- Piburn, M., & Sawada, D. (2000). Reformed teaching observation protocol (RTOP) reference manual. Technical Report. No. IN00-32. Arizona State University. <https://files.eric.ed.gov/fulltext/ED447205.pdf>

- Pöysä, S., Vasalampi, K., Muotka, J., Lerkkanen, M. K., Poikkeus, A. M., & Nurmi, J. E. (2019). Teacher-student interaction and lower secondary school students' situational engagement. *British Journal of Educational Psychology, 89*(2), 374–392.
- Roth, K. J., Druker, S. L., Garnier, H. E., Lemmens, M., Chen, C., Kawanaka, T., Rasmussen, D., Trubacova, S., Warvi, D., Okamoto, Y., Gonzales, P., Stigler, J., & Gallimore, R. (2006). *Teaching science in five countries: Results from the TIMSS 1999 video study* (NCES 2006-11). U.S. Department of Education, National Center for Education Statistics/U.S. Government Printing Office.
- Sammons, P., Kington, A., Lindorff-Vijayendran, A., & Ortega, L. (2014). *Inspiring teachers: Perspectives and practices*. CfBT Education Trust.
- Sammons, P., Lindorff, A. M., Ortega, L., & Kington, A. (2016). Inspiring teaching: Learning from exemplary practitioners. *Journal of Professional Capital and Community, 1*(2), 124–144.
- Schultz, S. E., & Pecheone, R. L. (2014). Assessing quality teaching in science. In *Designing teacher evaluation systems: New guidance from the measures of effective teaching project* (pp. 444–492). Jossey-Bass.
- Stigler, J. W., Gonzales, P., Kawanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS videotape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States*. From <https://nces.ed.gov/pubs99/1999074.pdf>
- Stigler, J. W., Gallimore, R., & Hiebert, J. (2000). Using video surveys to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS video studies. *Educational Psychologist, 35*(2), 87–10.
- Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (5th ed.). Pearson.
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273–1296.
- Taut, S., Jiménez, D., Puente-Duran, S., Palacios, D., Godoy, M. I., & Manzi, J. (2019). Evaluating the quality of teaching: Can there be valid differentiation in the middle of the performance distribution? *School Effectiveness and School Improvement, 30*(3), 328–348.
- van de Grift, W. (2007). Quality of teaching in four European countries. *Educational Research, 49*, 127–152.
- van de Grift, W. (2014). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement, 25*(3), 295–311.
- Virtanen, T. E., Lerkkanen, M. K., Poikkeus, A. M., & Kuorelahti, M. (2015). The relationship between classroom quality and students' engagement in secondary school. *Educational Psychology, 35*(8), 963–983.
- Walkington, C., & Marder, M. (2014). Classroom observation and value-added models give complementary information about quality of mathematics teaching. In T. Kane, K. Kerr, & R. Pianta (Eds.), *Designing teacher evaluation systems: New guidance from the measuring effective teaching project* (pp. 234–277). Wiley.
- Walkington, C., & Marder, M. (2018). Using the UTeach observation protocol (UTOP) to understand the quality of mathematics instruction. *ZDM Mathematics Education, 50*, 507–519.
- Wallace, T. L., Parr, A. K., & Correnti, R. J. (2020). Assessing teachers' classroom management competency: A case study of the classroom assessment scoring system—secondary. *Journal of Psychoeducational Assessment, 38*(4), 475–492.
- World Bank. (2019). *TEACH manual*. World Bank. https://www.worldbank.org/en/topic/education/brief/teachhelping-countries-track-and-improve-teaching-quality?cid=EXT_WBEmailShare_EXT
- Zhao, Y., & Ko, J. (2022, Apr). *Classroom orchestration as a potent stimulator of dialogic teaching and learner engagement in higher vocational classrooms*. [Paper presentation] American Educational Research Association (AERA 2022) Annual Meeting (Virtual), San Diego, California.

Dr. James Ko is an Associate Professor at the Department of Policy Leadership and Co-Director of the Joseph Lau Luen Hung Charitable Trust Asia Pacific Centre for Leadership and Change at the Education University of Hong Kong. Before his doctoral study, James was an EFL teacher for about 20 years and led two functional teams in a secondary school for 10 years. He is a recurrent grantee of the RGC and UGC grants and the principal investigator of 23 projects, collaborating with local academics and overseas researchers on 40 projects. He has supervised 14 doctoral students with 8 completed.

Zhijun Chen is a Doctoral Researcher at the Department of Education at the University of Bath (UK). She holds an MSc in Psychology from the University of St Andrews (UK). She also works as a research assistant with Dr James Ko at the Education University of Hong Kong, focusing on teaching quality and teaching assessment in multiple cultures. Her research interests include educational effectiveness, large-scale international assessments (PISA, TIMSS, PIRLS, ERCE, etc.), education inequality, and classroom observation.

Celia Lei is a lecturer at Shaoyang University in Mainland China and a doctoral candidate at the Education University of Hong Kong, where she graduated with an MEd degree and worked as a research assistant at the Department of Educational Policy and Leadership. She published on teaching quality of public schools in an inland province in Mainland China and has extensive research experience using various classroom observation instruments. Her research interests include dialogic teaching, teaching quality, and metacognitive teaching.

Ridwan Maulana is an associate professor at the Department of Teacher Education, University of Groningen, the Netherlands. His major research interests include teaching and teacher education, factors influencing effective teaching, methods associated with the measurement of teaching, longitudinal research, cross-country comparisons, effects of teaching behaviour on students' motivation and engagement, and teacher professional development. He has been involved in various teacher professional development projects including the Dutch induction programme and school-university-based partnership. He is currently a project leader of an international project on teaching quality involving countries from Europe, Asia, Africa, Australia, and America. He is a European Editor of Learning Environments Research journal, a SIG leader of Learning Environments of American Educational Research Association, and chair of the Ethics Commission of the Teacher Education.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Part III

Effective Teaching: Comparison Across Countries

Part III Overview

This part presents seven chapters that focus on cross-national comparisons of teaching behaviours, pedagogies, and student-teacher relationships using different instruments and the factors influencing differentiated instruction across countries.

Chapter 17 presents a longitudinal observation study involving more than 3000 teachers across 5 countries. Results show that within-teachers differences are consistently large across countries and that the amount of between-schools and between-teachers differences vary depending on the country and the teaching behaviour domain. Implications for practice and policy are discussed.

The review presented in Chap. 18 evaluates a teacher-student interaction framework in different contexts to examine the effect of interaction quality on student outcomes. The applicability of the framework is discussed and suggestions are provided for educational policy and future research.

An empirical comparison study of affective student-teacher relationships in The Netherlands and China is presented in Chap. 19. The results reveal that closeness may be more relevant for Chinese student engagement that expected and conflict seems to be equally harmful in both cultures. The authors conclude that developing relationship-focused interventions for Chinese teachers and students seems important.

The relationship between student perceptions of teaching behaviour and engagement across six countries is presented in Chap. 20 using data of more than 35 000 students collected in six countries. Results show that the quality of perceived teaching behaviour is strongly and positively related to student engagement in all the six involved countries. Student background does not play a significant role in this relationship. Implications for research and practice are discussed.

Chapter 21 reviews the literature on the play-based pedagogies in mainland China, Hong Kong, Singapore and Japan. The authors conclude that the existing publications in English academic journals, reflect traditional Asian values and deep-rooted beliefs regarding Early Childhood education, where play is seen as a rather

unimportant activity. They call for more extensive, rigorous, and locally situated play impact studies.



The reviews study presented in Chap. 22 on effective interpersonal relationships between teachers and students reveals that the communion and agency interpersonal dimensions are related to cognitive and affective outcomes. The agency dimension is more related to cognitive outcomes and the communion dimension is more related to affective outcomes.

The last chapter of this part, Chap. 23, presents a theoretical and empirical exploration of the influence of teacher characteristics and contextual factors on differentiated instruction across six countries.

Chapter 17

Secondary Education Teachers' Effective Teaching Behaviour Across Five Countries: Does it Change Over Time?



Ridwan Maulana , Amanda Maraschin Bruscato, Michelle Helms-Lorenz , Yulia Irnidayanti, Thelma de Jager, Ulziisaikhan Galindev, Amarjargal Adiyasuren, Abid Shahzad, Nurul Fadhilah, Seyeoung Chun, Okhwa Lee, Thys Coetzee, and Peter Moorero

Abstract Over the last decade, a limited number of studies have documented changes in effective teaching behaviour in secondary education over time. However, the studies are rather fragmented and heterogeneous in terms of measurements, contexts, and time intervals.

This study aims to investigate changes in secondary school teachers' teaching behaviour over time, by using a uniform observation instrument in five contrasting national contexts. The study focuses on the examination of inter- and intra-individual differences in teachers' effective teaching behaviour across Indonesia, Mongolia, Pakistan, South Africa, and the Netherlands. A total of 3158 teachers across the five countries participated in this study. Their classroom lessons were observed by trained observers in the natural classroom setting longitudinally using a uniform observation measure called International Comparative Analysis of Learning and Teaching (ICALT). Results show that, in general, between-schools, between-teachers, and within-teacher differences are visible, with some degree of variations in proportion depending on the country and the type of teaching behaviour. Within-teacher differences are consistently large across countries. This provides evidence regarding the dynamic characteristics (i.e., change) of teaching behaviour cross-nationally. Implications for research and practice are discussed.

R. Maulana (✉) · M. Helms-Lorenz
Department of Teacher Education, University of Groningen, Groningen, The Netherlands
e-mail: r.maulana@rug.nl; m.helms-lorenz@rug.nl

A. M. Bruscato
Faculty of Human and Social Sciences, University of Algarve, Faro, Portugal

Y. Irnidayanti
Department of Biology and Biology Education, Faculty of Mathematics and Science, State University of Jakarta, Jakarta, Indonesia

T. de Jager · T. Coetzee
The Department of Educational Foundation, Tshwane University of Technology,
Pretoria, South Africa
e-mail: DeJagerT@tut.ac.za

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_17

Keywords Effective teaching behaviour · Cross-national study · Longitudinal study · Multilevel growth curve modelling · Secondary education

1 Introduction

Research on classroom practice and teacher effectiveness has shown that classroom factors contribute more variance to explain student attainment than school factors. Within the classroom factors, what the teacher does in the classroom matters the most (Muijs et al., 2014; Coe et al., 2014). In teacher effectiveness research, attempts to uncover effective teaching behaviour have motivated scholars to investigate behaviours using observation instruments (Muijs et al., 2018). Particularly, effective teaching behaviour, which is the focus of the present study, has grown to become a central theme internationally, which is reflected in terms of the existence of various observation instruments tapping teachers' classroom behaviour (e.g., Danielson, 2013; Pianta et al., 2008; Reynolds et al., 2002; van de Grift et al., 2014).

The positive effects of various effective teaching behaviour on student outcomes have been well-documented in the literature (e.g., Maulana et al., 2017; Pianta et al., 2008, Stroet et al., 2015). However, little is known about whether and how teachers change their teaching behaviour over time, taking into account a uniform measure across various national contexts. This knowledge is important for at least two reasons. Firstly, it can add to the knowledge base regarding the dynamic characteristics of teaching behaviour across various contexts (i.e., trend specificity versus generality). Secondly, it provides insights into the temporal aspect of change over time across contexts (i.e., inter-individual versus intra-individual variability), indicating dynamic and critical points of changes in teaching effectiveness over time.

U. Galindev

The Department of Educational Administration, Mongolian National University of Education, Ulaanbaatar, Mongolia
e-mail: ulziisaikhan@msue.edu.mn

A. Adiyasuren

Mongolian National University of Education, Ulaanbaatar, Mongolia

A. Shahzad

Department of Education, The Islamia University of Bahawalpur, Bahawalpur, Pakistan

N. Fadhilah

Department of Biostatistics and Population Studies, Faculty of Public Health, Universitas Indonesia, Depok City, Indonesia

S. Chun

Department of Education, Chungnam National University, Daejeon, South Korea

O. Lee

Department of Education, Chungbuk National University, Cheongju, South Korea

P. Moorer

University of Groningen, Groningen, The Netherlands

Research on teaching quality started in the early sixties, although most of the past studies have focused on very specific contexts, with a small number of teachers and schools (van de Grift, 2014). Moreover, studies in which changes in teachers' observed classroom quality have been investigated did not involve multiple countries. Over the last decade, a limited number of studies have documented changes in effective teaching behaviour in secondary education over time (e.g., Mainhard et al., 2011; Malmberg et al., 2010; Maulana, 2012). However, the studies are rather fragmented and heterogeneous in terms of measurements, contexts, and time intervals.

Regarding the measurement, past studies typically investigated teaching behaviour using various (observation) instruments. The heterogeneity of the instruments used poses challenges for comparing teaching behaviour across studies. Although the teaching behaviour construct is used across studies, its operationalization often varies with a relatively moderate degree of overlap (Maulana & Helms-Lorenz, 2016). Even when the same measure is used, the equivalence of the measure across contexts cannot be fully guaranteed (Maulana et al., 2020a, b; Muijs et al., 2018). Regarding contexts of studies, the focus has been mainly on a single context within a single country (e.g., Mainhard et al., 2011; Malmberg et al., 2010; Opdenakker et al., 2012). Although single-context studies on teaching behaviour add to the knowledge base from certain contexts which can serve as building blocks for potentially higher level knowledge (e.g., generic teaching behaviour), the transferability of the findings to other contexts is limited. Regarding the time intervals, past studies vary in their investigation from between-week (e.g., Mainhard et al., 2011), between-month (e.g., Maulana et al., 2012; Opdenakker et al., 2012) to between-year (e.g., Malmberg et al., 2010; Maulana et al., 2015) periods of time. The time intervals discrepancy between studies limits the comparability of changes between contexts because changes that happen during shorter periods (e.g., weeks) cannot be compared with changes during longer periods (e.g., months or years) due to differences in the personal and contextual factors operating between different time-spans.

The aim of the current study is to investigate changes in secondary school teachers' teaching behaviour over time, by using a uniform observation instrument in five contrasting national contexts. Particularly, we focus on the examination of inter- and intra-individual differences in teachers' effective teaching behaviour across countries including Indonesia, Mongolia, Pakistan, South Africa, and the Netherlands.

2 Theoretical Framework

2.1 *Effective Teaching Behaviour*

Teaching behaviour is a multidimensional concept (Shuell, 1996) addressed in the literature on teacher effectiveness, learning environments, and motivation. It can be measured with different methods and instruments, such as observations, interviews,

and surveys, and it has been proved to be essential to students' learning outcomes (Brophy, 1986).

It has already been shown that observed teacher support is a strong predictor of student engagement and achievement (Roorda et al., 2011). However, there is a preference, especially in secondary education, to use surveys instead of classroom observations to analyse teaching quality and student engagement (Virtanen et al., 2015). Although surveys have better cost and time effectiveness, classroom observations have been shown to be more objective (Worthen et al., 1997). Thus, the present study focuses on the effectiveness of observable behaviours in a classroom during regular lessons.

There are many observation instruments to investigate teaching behaviour (for a review, see Sandilos et al., 2019). Although the instruments on effective teaching have differences, they also share some similarities (van de Grift et al., 2017). Although teaching quality is a broad concept, it usually includes an interpersonal component, an instructional component, and a structural component (Maulana, 2012).

For this study, the International Comparative Analysis of Learning and Teaching (ICALT; van de Grift, 2014) observation instrument was used to measure effective teaching behaviour, since it has been validated for use in secondary education in multiple countries (Maulana et al., 2020a). It is grounded in evidence-based teacher effectiveness research and has six observable domains comparing interpersonal, instructional, and structural components: creating a safe and stimulating learning climate (Learning climate), providing efficient classroom management (Classroom management), displaying clarity of instruction (Clarity of Instruction), activating teaching (Activating teaching), adapting instruction to students' learning needs (Differentiated instruction), and teaching students learning strategies (Teaching learning strategies).

2.2 Inter-personal and Intra-personal Variability in Teaching Behaviour

Patterns of change over time can be distinguished between interpersonal and intrapersonal differences. Interpersonal differences in change over time refers to the variation of the shape and pacing of change from one individual to the other, whereas intrapersonal change over time indicates variations of ups and downs across moments or situations (Malmberg et al., 2010). Fuller's (1970) stage theory of concerns relates to teacher interpersonal differences in their development across the career trajectory. The theory postulates that teachers' concerns can be distinguished into three stages: concerns about self (first stage), concerns about tasks (second stage), and concerns about impact on students (third stage (Fuller, 1970). Research shows that self-related concerns decline during the pre-service period, and beginning in-service teachers reported more student-impact concerns and fewer self-related concerns. Intrapersonal differences in change has been argued to be related

to teachers' personal characteristics and contextual factors. Particularly, teacher flexibility, adaptation to, and coping with situational demands can be viewed as factors related to intrapersonal variability.

Studies investigating inter-personal and intra-personal variability in teaching behaviour using observation instruments across countries are scarce.¹ Past studies typically included single or two country contexts only. For example, studies by and Mayer and Seidel et al. in Germany both cited in Kunter and Baumert (2006) found stability of teaching behaviour over several weeks. Malmberg et al. (2010) found larger intra-personal than inter-personal variability in teaching behaviour among secondary school teachers in England across three years. Similarly, Maulana et al. (2013) found larger intra-personal than inter-personal variability in teaching behaviour among secondary school teachers in Indonesia and the Netherlands across the school year. Based on these studies, there seems to be a general tendency that large intra-personal differences in teaching behaviour tend to be detected in studies employing longer time lags (e.g., months to years) than shorter time lags (e.g., weeks). Although the multi-country perspectives on these studies are limited, the current evidence seems to point out the relevance of this trend across countries. Factors that can explain intra-personal variability (between lessons) in teaching behaviour are unclear. Some potential causes like lesson materials and content characteristics, teaching modes, and developing interpersonal relationships between teachers and students may be worth investigating in future research.

2.3 Differences and Changes in Teaching Behaviour Across Countries

Studies on changes in effective teaching behaviour (using observation instruments) typically focus on one or two contexts or countries. In the Netherlands, for example, Stroet et al. (2015) observed 20 math teachers and assessed 489 students' motivation in four moments during the first year of secondary education. They analysed videotaped lessons using a rating sheet to assess need-supportive teaching from the perspective of the self-determination theory and found declining trends for the teachers' levels of autonomy support and involvement, but an upward trend for structure. Also in the Netherlands, 1208 secondary students were asked to analyse the behaviour of 48 teachers and their classroom social climate (Mainhard et al., 2011). Teachers' behaviour had a direct correlation with the classroom social climate during the current lesson and in the lesson a week later in terms of teachers' proximity (Affiliation), but not in terms of teachers' influence (Control). Classroom social climate did not change much from their initial status, and the development of teacher Affiliation over time was related to its perception in the first lesson. It was

¹Inter-personal and intra-personal variability in teaching behaviour measured by other than observation (e.g., student and teacher questionnaire) is beyond the focus of this study.

more likely to decline in classrooms that already started with lower levels of Affiliation (Mainhard et al., 2011).

Using the Classroom Assessment Scoring System observation instrument (CLASS-S), Malmberg et al. (2010) measured the teaching quality of 17 secondary school teachers in England during their initial postgraduate preservice teacher education year and their first two years of teaching. They found a linear increase in classroom organization, an increase followed by a decrease in emotional support, and no change in instructional support over time. Since the studies mentioned used different observation instruments and time-spans to analyse teaching behaviour, their results are difficult to compare. While they focused on single countries, an example of an ambitious project that analyses effective teaching behaviour in 20 countries is the International System for Teacher Observation and Feedback (ISTOF) (Reynolds et al., 2002). However, for cross-country comparisons, its factor invariance across contexts remains unknown (Muijs et al., 2018). Some studies compared teaching quality between two countries, such as between the Netherlands and Indonesia (Maulana, 2012), between the Netherlands and South Korea (van de Grift et al., 2017; Maulana et al., 2020b), and between South Korea and Mongolia (Chun et al., 2020).

According to Stigler et al. (1999), the International Mathematics and Science Study (TIMSS) 1995 Video Study was the first research to use videotaped lessons to investigate teaching across countries. However, this large-scale study is not longitudinal. In 1999, they compared the teaching practices of mathematics and science in eighth-graders in Australia, the Czech Republic, Hong Kong SAR, Japan, the Netherlands, Switzerland, and the United States (Givvin et al., 2005). Although many lessons were recorded during a school year, each teacher was observed only once. The project aimed at identifying national patterns of teaching, but without a focus on changes between teachers over-time. In Chile, 51 secondary math teachers had their lessons recorded twice in one year (Bruns et al., 2016). The videos were analysed using two different observation instruments: CLASS-S and the Stanford Research Institute Classroom Observation System (Stallings, 1977; Stallings & Mohlman, 1988). The results showed that the quality of instruction and the emotional support were better captured by CLASS-S. However, the Stallings instrument seemed more suitable for larger-scale studies. The Chilean results were then compared with results from other six Latin American countries, showing that Chilean math teachers managed to keep their students more engaged in the lessons.

While the study above focused on Latin American countries, Maulana et al. (2020a) observed and analysed secondary teacher's behaviour in the Netherlands, South Korea, Indonesia, South Africa, Hong Kong - China, and Pakistan. The results show that South Korean teachers usually scored higher than the other countries included, while Indonesian teachers were generally rated lower by observers. They also found that, for most of the countries, differentiated instruction was rated the lowest. In general, there is an indication that interpersonal and intrapersonal variability in teaching behaviour seems to be visible in various country contexts.

2.4 Contexts of the Present Study

The present study focuses on teaching behaviour in secondary education in five contrasting countries with different educational systems. The countries' context information is briefly provided below.

The Netherlands Academic tracking is employed in Dutch secondary education. Students perform above average in international comparisons (OECD, 2018). Teachers do not have an above average professional status, although the professional quality is generally high (OECD, 2016a). Recent research indicates that Dutch teachers are generally skillful in teaching behaviour related to classroom climate, classroom management, instructional clarity, and activating teaching. However, their skills in differentiated instruction and teaching learning strategy are relatively low (Maulana et al., 2020).

Indonesia The Indonesian educational system has been among the lower-performing countries in international comparisons (OECD, 2016b). This trend has been argued to be caused by low quality teaching. Recent research partially confirms this argument (Maulana et al., 2020). Although culturally the teaching profession is typically highly respected, the profession is not viewed as a high-status profession (Maulana et al., 2011).

South Africa The South African educational system is considered as one of the lowest performers in the world (Baller et al., 2016; Mullis et al., 2017). Some of the country's challenges include: the English second language instruction barrier, insufficient subject knowledge of teachers, lack of accountability of teachers, frequent absenteeism of teachers from classes, and socio-economic status of most students (Mbiti, 2016). As a consequence of the mentioned poor quality education indicators, other problems including unemployment, poverty, and inequality may increase (Van der Berg & Hofmeyr, 2017).

Pakistan Pakistan has the third largest adult illiteracy in the world and almost half of the young rural women never even get the chance to go to school (UNESCO, 2015). The quality of initial teacher education is below the international standards (UNESCO, 2006) and the country ranks 113th out of 120 countries on the educational performance index (UNESCO, 2015). The teaching profession is considered a low status activity (Khan, 2019).

Mongolia Mongolia shifted from one of the worst educational systems in Central Asia in early 1990's to one of the region's top performers (UNESCO, 2020). The country is planning to participate in PISA for the first time in 2021. To date, comparative education data from Mongolia for international large-scale assessments is not available. Teacher qualification examination was introduced in 2014 for new graduates in order to be qualified to become a novice teacher; however, the regulation was withdrawn in 2018. Policy review suggests that the qualification

examination should be recovered back legally (UNESCO, 2020). The teaching profession is regarded as a low paid profession and there is no competition for teacher recruitment.

2.5 Research Questions

The current study focuses on the examination of inter- and intra-individual differences in teachers' effective teaching behaviour across countries including Indonesia, Mongolia, Pakistan, South Africa, and the Netherlands. The research questions are as follows:

1. How variable is effective teaching behaviour over time, between teachers, and between schools across the five countries?
2. Does effective teaching behaviour change over time across the countries?
3. Are there differences between individual teachers with regard to change over time across the five countries?

3 Method

3.1 Sample and Procedure

The data was drawn from a large longitudinal research project on effective teaching behaviour involving over 16 countries across the globe. For the present study, available longitudinal data from Indonesia, Mongolia, Pakistan, South Africa, and the Netherlands are included. Data from South Korea is also available. Unfortunately, it was difficult to measure teachers over time partly due to the teacher rotation policy, so the Korean data was excluded. Other participating countries did not collect longitudinal data. The initial plan was to collect longitudinal data in the five countries employing similar time intervals: three measurement moments, once a year. Hence, between-years change in teaching behaviour was focused on. However, not all countries made it to meet the initial requirement due to highly challenging circumstances (e.g., financial, bureaucracy, resources, and field issues). Hence, a realistic approach to data collection was applied.

In the Netherlands, data were collected in four measurement moments across three school years (twice in year 1, once in the subsequent years). In South Africa and Indonesia, data were collected once a year for two and three school years respectively. In Pakistan and Mongolia, data were collected in two and three measurement moments respectively, based on a semester interval. Natural classroom observations took place between the school year of 2015 and 2019. A simple random sampling procedure was planned. However, this design was not implemented successfully. Teachers participated on a voluntary basis.

Table 17.1 Sample demographics

Country	N. School	N. Teacher	N. Measurement	Female (%)	Science subject (%)	Years of experience	Public school (%)
Indonesia	27	454	3	62	42	16.61 (SD 9.78)	78
Mongolia	52	375	2	85	47	11.06 (SD 8.66)	91
Pakistan	18	336	2	43	48	6.30 (SD 5.14)	100
South Africa	35	316	2	49	42	N.A.	98
The Netherlands	350	1677	4	57	33	1.92 (SD 2.32)	100

The present study included 454 teachers from 27 schools in Indonesia, 375 teachers from 52 schools in Mongolia, 336 teachers from 18 schools in Pakistan, 316 teachers from 35 schools in South Africa, and 1677 teachers from 350 schools in the Netherlands (see Table 17.1 for more demographic information).

3.2 Measures

Teaching behaviour was measured using the International Comparative Analysis of Learning and Teaching observation instrument (van de Grift et al., 2014). The instrument consists of 32 high inferential observable teaching behaviours, accompanied with 120 low inferential observable teaching indicators. The high inference items represent the six domains of teaching behaviour: safe and stimulating learning climate (4 items), efficient classroom management (4 items), clarity of instruction (7 items), activating teaching (7 items), differentiated instruction (4 items) and teaching learning strategies (6 items). Previous research has confirmed the six-factor structure of observed teaching behaviour in the five countries (Maulana et al., 2020). Observers rated the items on a four-point scale, ranging from 1 ('mostly weak') to 4 ('strong').

3.3 Translation and Back-Translation

The instrument was originally developed in Dutch. The original English version of the instrument was used as the source language for the translation and back-translation procedure. The target language of translation includes Indonesian and Mongolian. In South Africa and Pakistan, English is used as language of instruction. Hence, the English version was used in these countries. The guidelines of the

International Test Commission (Hambleton, 1994) were followed. The process involved two highly knowledgeable researchers concerning the instrument and the theoretical framework underlying the instrument and two university professors proficient in both English and the target languages. Upon the completion of the procedure, issues and discrepancies were discussed thoroughly and resolved subsequently by the team. The national expert team checked and confirmed the relevance of the six domains of teaching behaviour in their own national contexts, providing evidence of face validity.

3.4 Observer Training

In the five countries, the onsite observer training for using the ICALT observation instrument was conducted applying identical standards, structure, and procedure. Two expert trainers led the training in the five countries, assisted by local trainers who were already trained earlier. Due to challenging circumstances, the training in Pakistan was conducted online using a digital platform.

The training consisted of 3 phases: preparation, implementation, and evaluation. In the first phase, the trainees studied the theoretical framework underlying the instrument and the content of the instrument thoroughly. In the second phase, the trainees attended a full day training covering the presentation and discussion about the instrument as well as how to rate indicators of teaching behaviour using the applied scoring rules. Subsequently, they practiced scoring two video-taped lessons using the instrument. The consensus level of 70% within the group and between the group and the expert norm was set as a cut-off criteria. Discussion to resolve significant differences and improve consensus were conducted subsequently. Finally, the third phase involved the investigation of rating patterns and significant deviations from the average pattern. A small number of observers who deviated from the average were followed up and extra guidance was given to this group prior to conducting the observation in the natural classroom settings. Observers failing to meet the minimum consensus were not invited to conduct observations. The consensus level was found to be satisfactory, ranging from 63% (Pakistan) to 88% (South Africa). A slightly lower consensus percentage for Pakistan might be caused by the online training approach because onsite training was not possible at that time.

3.5 Analysis Technique

The data of the present study were systematically structured in a hierarchical order (i.e., measurement moment, teacher, school). Multilevel modelling is appropriate for analysing this type of data (Snijders & Bosker, 2012). When the data are ordered hierarchically and longitudinally, multilevel growth curve modelling (MLGCM), is the most appropriate approach. With this method, not only the hierarchical structure

of the data is taken into account, but also the multiple measurements over time and predictor variables. MLGCM is an extension of mixed-effect regression model (MRM) applied to multilevel and longitudinal data (Rasbash et al., 2014).

The first research question was related to the relative proportions of explained variance across levels. To answer this question, we performed MLGCM and interpreted results based on the baseline model (Model 0). The second research question is related to the shape of change over time. We included fixed effects of time (linear, quadratic) to the model (Model 1). The quadratic term was only included when there were more than two measurement moments in the data. The third research question is related to the extent to which individual differences in change could be observed. We added random effects of time (linear) to the model (Model 2) and a covariance term at the teacher level (i.e., whether the time slopes vary across teachers). The modelling was done using a stepwise procedure and separately for each domain of effective teaching behaviour and for each country data. Significant levels at $p < .05$ were retained. The fixed effects in the model were tested by using t-ratio coefficients for a significant effect of a variable, and the random effects were tested by comparing two competing models (Snijders & Bosker, 2012).

4 Results

4.1 Variability of Effective Teaching Behaviour

Based on the MLGCM baseline model (see Table 17.2, also Appendix A), we found relatively more variability within teacher over time (41–99%) than between teachers and between schools for all domains of effective teaching behaviour in all countries, except for Differentiated Instruction in Indonesia. This means that, in general, teaching behaviour is not stable over time across countries. In Indonesia, between schools variability in Differentiated Instruction was larger (71%) than between teachers (5%) and within teachers over time (54%). This indicates that schools in Indonesia differ greatly in the quality of Differentiated Instruction. Although the amount of within teachers variability was generally very large, the magnitude of the variance differed across countries, ranging from 41–54% in Indonesia, 58–67% in the Netherlands, 59–86% in Mongolia, 70–80% in Pakistan, and 88–99% in south Africa.

In the Netherlands, between teachers variability was generally larger (17–23%) than between schools (Netherlands: 10–16%). There is an exception for Teaching Learning Strategy in the Netherlands, in which between teachers variability was relatively smaller (17%) than between schools (23%). This means that, in general, differences between schools and between teachers in teaching behaviour are visible.

In Indonesia, Mongolia, Pakistan, and South Africa, between teacher variability in effective teaching behaviour was generally smaller (Indonesia: 5–24%, Mongolia: 12–18%, Pakistan: <1%, South Africa: <1%) than between schools (Indonesia:

Table 17.2 Proportion of variance across school, teacher, and measurement moment levels

Country	Level	CLM	ORG	CLR	ACT	DIF	TLS
The Netherlands	School	14.69	11.41	10.22	16.75	14.53	23.12
	Teacher	21.68	22.82	23.53	18.23	18.80	16.27
	Moment	63.64	65.77	66.25	65.02	66.67	60.62
Indonesia	School	20.85	35.15	35.53	31.68	39.46	35.81
	Teacher	25.68	19.17	23.92	21.21	5.61	16.59
	Moment	53.47	45.68	40.55	47.11	54.93	47.60
South Africa	School	17.37	9.6	10.75	11.21	10.16	9.20
	Teacher	0.001	0.001	0.001	0.001	0.001	0.001
	Moment	82.62	90.35	89.24	88.78	89.83	90.8
Mongolia	School	22.16	18.58	17.51	20.90	25.33	25.41
	Teacher	12.05	16.79	18.19	15.99	15.39	13.75
	Moment	85.74	64.63	64.29	63.11	59.27	60.83
Pakistan	School	20.35	25.83	26.44	29.56	26.31	26.52
	Teacher	0.001	0.001	0.001	0.001	0.001	0.001
	Moment	79.65	74.17	73.56	70.44	73.69	73.48

Note: *CLM* Climate, *ORG* Classroom management, *CLR* Clarity of instruction, *ACT* Activating teaching, *DIF* Differentiated teaching, *TLS* Teaching learning strategy

22–71%, Mongolia: 18–25%, Pakistan: 20–30%). However, there are also some exceptions. In Indonesia, between-teachers variability in Learning Climate was larger (24%) than between schools (22%). In Mongolia, between teachers and between schools variability was about the same (18%) for Clarity of Instruction. In Pakistan and South Africa, between schools variability was generally moderate (South Africa, 10–17%) to large (Pakistan, 20–30%). This means that, in general, differences between schools and between teachers in teaching behaviour are visible in these countries. The negligible variability at the teacher level in Pakistan and South Africa was also visible (<1%) (see Table 17.2, also Appendix A). This means that in these two countries, between teacher differences in teaching behaviour in general are not visible. This may suggest that the quality of teaching behaviour of teachers in these two countries is homogeneous.

4.2 Change in Effective Teaching Behaviour over Time

Based on the MLGCM fixed time effect (see Fig. 17.1,² also Appendix A), we found differences in the pattern of change over time in the five countries. In Pakistan and South Africa, only two measurement moments are available (only linear trend can be estimated). The change of effective teaching behaviour in these two countries showed a linear increase from moment 1 to moment 2 ($p < 0.05$). In the Netherlands

²Some lines may visually look like straight lines due to the scaling of the graph.

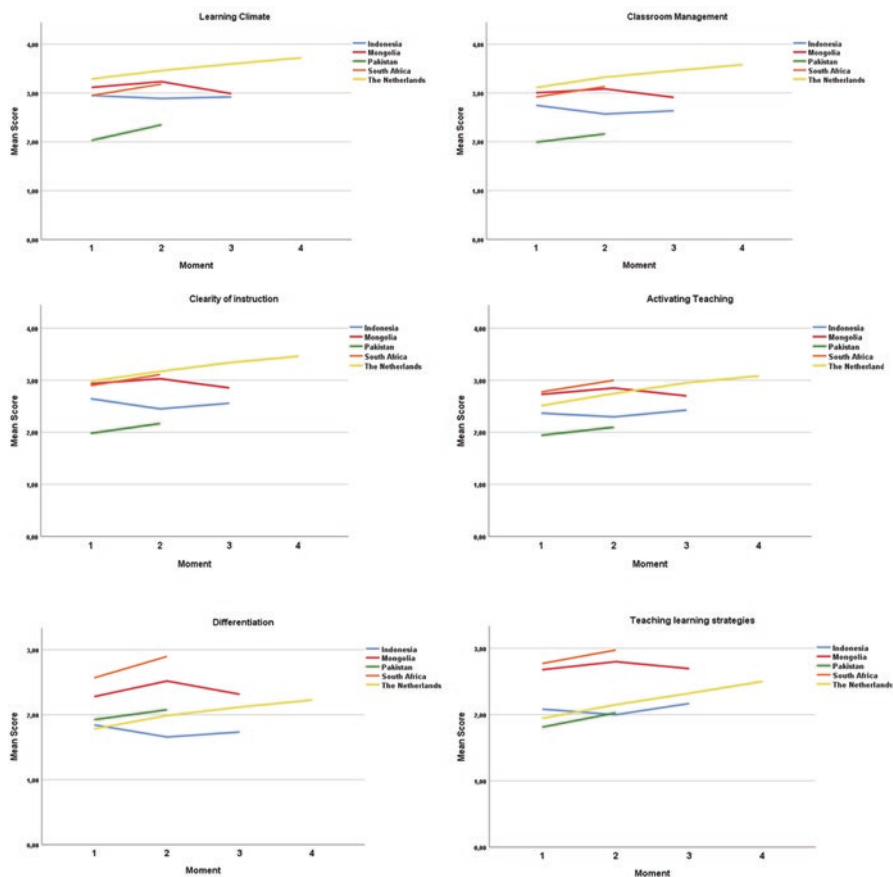


Fig. 17.1 Changes in teaching behaviour over time across countries

and Mongolia, the change in effective teaching behaviour exhibited curvilinear, inverted U-shaped like, patterns ($p < 0.05$). However, the inverted U-shaped like pattern in Mongolia was steeper compared to the Netherlands. In the Netherlands, effective teaching behaviour generally increased significantly from moment 1 to moment 4. The increase was steeper from moment 1 to moment 2, then it decelerated slightly between moment 2 and moment 4. In Mongolia, effective teaching behaviour also increased from moment 1 to moment 2, and it decreased between moment 2 and moment 3 subsequently. In general, the pattern of change is consistent for all domains of teaching behaviour across the five countries.

In Indonesia, the change of effective teaching behaviour was best represented by a curvilinear, U-shaped pattern ($p < 0.05$), except for learning climate ($p > 0.05$). For the five domains of effective teaching behaviour, the change was marked by a decrease from moment 1 to moment 2, then it continued to increase from moment 2 to moment 3.

4.3 *Individual Differences in Change Over Time*

Based on the MLGM random effect of time and the covariance terms between the intercepts and the slopes at the teacher level (see Appendix A), we found negative covariance coefficients between intercepts and slopes for all six teaching behaviour domains ($p < 0.05$). This trend is consistent for teaching behaviour domains across the five countries. This means that, in general, teachers who started off lower in effective teaching behaviour during the first measurements showed steeper increases over time compared to those who started off higher at the end of the measurements.

5 **Conclusions and Discussion**

The present study aimed to investigate inter-individual (between teachers) and intra-individual (change) differences in effective teaching behaviour over time across five contrasting countries including the Netherlands, South Africa, Indonesia, Mongolia, and Pakistan. We focused the investigation on: (1) variability across levels (school, teacher, time), (2) general and pattern of change over time, and (3) individual differences in change over time.

We found generally larger intra-personal than inter-personal differences in teaching behaviour across the five countries. This implies that, in general, greater differences in the six domains of teaching behaviour are attributed to within teacher practices over time, irrespective of the countries. In the teaching context, intrapersonal variability can be viewed in multiple ways. It can be related to teacher flexibility in modifying their behaviour in line with the daily classroom dynamics. It can also be perceived as teacher adaptation to the classroom dynamic situation, as well as a way of coping with situational demands.

The magnitude of within-teacher variability was largest in South Africa (83–91%), followed by Pakistan (70–80%), Mongolia (59–86%), the Netherlands (58–67%), and Indonesia (41–54%) respectively. The differences in the magnitudes of intra-personal variability across countries may be related, at least to some degree, to the differences in the measurement intervals. The results indicate that larger intra-personal variabilities seem to be more evident in countries with shorter measurement intervals (Pakistan and Mongolia) compared with longer measurement intervals (Indonesia and the Netherlands). This indicates that teaching behaviour may be more dynamic within the school year compared to between school years. Nevertheless, this trend does not seem to apply to South Africa.

Our findings may also suggest that in general, teaching behaviour of teachers in South Africa, Pakistan, and Mongolia seems to be more prone to changes over time due to the contextual differences where they teach, requiring them to employ greater flexibility in their teaching practice, adapt to daily situational dynamics, and cope with the dynamic of situational demands compared to that of teachers in the

Netherlands and Indonesia. This large variability may be related to language instruction barriers, insufficient subject knowledge, inadequate resources, and heavy workloads experienced by South African teachers (Lumadi, 2008; Mbiti, 2016), lack of resources, poor teacher quality, and lack of professional development opportunities experienced by Pakistani teachers (Ahmad et al., 2014), and dealing with vulnerable students and minorities, and children of herders experienced by Mongolian teachers (Steiner-Khamsi & Gerelmaa, 2008).

Interestingly, within teacher variability in differentiated instruction in Indonesia was smaller (54%) than between schools (71%), although the amount of intrapersonal variability remained reasonably large. Schools seemed to vary largely in differentiated instruction in Indonesia, implying that this teaching behaviour domain seems to operate as the between-school variable more strongly, which is quite unique compared to the other four countries. Although reasons for this finding remain unclear, it is possible that this finding is related to a large inequality between schools (and regions) in Indonesia (OECD & ADB, 2015), particularly in terms the opportunity of implementing differentiated instruction, which is seen as a contemporary trend in education (Maulana et al., 2020b; Smale-Jacobse et al., 2019).

An interview with an expert observer in Indonesia revealed that teachers in Indonesia tended to focus heavily on teaching the learning materials to achieve curriculum completeness (content knowledge focus) on time, paying little attention to students' diversity in the classroom. The teaching and learning process tended to be teacher-centred, giving little room for flexibility. In addition, the student recruitment system in Indonesian public schools is based on the ability ranking, whereby students with high rankings can enter public schools. This system is extremely competitive, and this group of students entering public schools typically have high academic motivation. This conditions forced teachers to focus on content knowledge heavily, and much less to pedagogical component like differentiated instruction. Teachers in many schools tended to employ a monotonous and one-size-fits-all approach. Only in some high-ranked schools the teachers paid a more attention to differentiation to a limited extent (N. Fadhillah, personal communication, Mei 21, 2021).

With respect to changes in effective teaching behaviour over time, the patterns of change differed depending on the country and measurement intervals. In the Netherlands, Mongolia, and Indonesia (≥ 3 measurements), the change followed a curvilinear trend. However, the direction of change differed. The inverted U-shaped like was evident in The Netherlands and Mongolia. However, the magnitude of change in these two countries also differed. From moment 1 to moment 2, teaching behaviour increased, then the increase continued over time with a slight deterioration in the Netherlands. In Mongolia, teaching behaviour decreased subsequently. These patterns of change might be related to the time interval when teaching behaviour was measured in the two countries. In the Netherlands, the measurement took place in four different moments across three years (between-year change). In Mongolia, the measurement took place in three different moments across three semesters.

The pattern of change in the Netherlands is in line with previous studies on supported beginning and in-service teachers, which suggested a steeper initial increase in teaching quality and a tendency to level off towards the end-of-time-span (Hebert

& Worthy, 2001; Maulana et al., 2015; Woolfolk-Hoy & Burke Spero, 2005). This change pattern might also be related to the characteristics of the Dutch sample, which was highly dominated by beginning teachers. The literature acknowledged that the first year of teaching is usually filled with optimism and commitment, although at the same time this period is often experienced as stressful (Hebert & Worthy, 2001; Woolfolk-Hoy & Burke Spero, 2005). Throughout the first years of professional practice, beginning teachers were found to be less democratic and more custodial over time (Hoy & Woolfolk, 1990).

The inverted U-shaped change in Mongolia might reflect the challenging nature of teaching in the Mongolian context, particularly in the beginning of the school year (first semester). In the second semester, teaching behaviour seemed to increase, but it continued to decrease again in the subsequent semester of the new school year. Again, this pattern of change may be related to the contextual and personal challenges faced by the Mongolian teachers (Steiner-Khamsi & Gerelmaa, 2008). In Indonesia, the change of effective teaching behaviour was best represented by a U-shaped like pattern ($p < 0.05$), except for learning climate ($p > 0.05$). The change was marked by a decrease from moment 1 to moment 2, then it continued to increase from moment 2 to moment 3. It is unclear what caused the decrease of teaching behaviour in the second year.

In Pakistan and South Africa, only two measurement moments are available so only linear changes can be estimated. A linear increase from moment 1 to moment 2 is visible ($p < 0.05$). In the Netherlands and Mongolia, the change in effective teaching behaviour exhibited inverted U-shaped patterns ($p < 0.05$). However, the inverted U-shaped pattern in Mongolia was steeper compared to the Netherlands. In the Netherlands, effective teaching behaviour increased significantly between moment 1 and moment 4. The increase was steeper from moment 1 to moment 2, and the increase was decelerated slightly between moment 2 and moment 4. In Mongolia, effective teaching behaviour also increased from moment 1 to moment 2, and it decreased between moment 2 to moment 3. In the five countries, in general, the pattern of change is consistent for all domains of teaching behaviour. The increase in teaching behaviour may be explained by increasing experience over time.

Regarding individual differences in change over time, it was found that in general, teachers who started off lower in effective teaching behaviour increased more over time than those who started off higher. This result was consistent across the five countries. This general individual pattern of change over time may represent a mastery effect (Malmberg et al., 2010).

5.1 Implications

The present study provides preliminary evidence of inter-individual and intra-individual differences in teaching behaviour across different national contexts. It supports the conceptualization of effective teaching behaviour as a dynamic characteristic that is subject to change over time, which may be universal irrespective of

the measurement moments and the national contexts. This finding implies that interventions to improve the quality of effective teaching behaviour should take into account inter-personal and intra-personal variability in teachers' teaching practices. Teacher professional development (PD) should be tailored in line with the unique and dynamic characteristics of teachers' teaching behaviour trajectory over time. This suggests that pedagogic and strategic content of PD programs should be made available considering the temporal, time-based, approach. The programs may include a semester-based intervention, an annual-based intervention, and regular between-years intervention. Such programs, if effectively tailored, may help to mitigate and even reverse the decline in teaching behaviour during certain schooling periods.

5.2 Limitations and Future Directions

The present study is subject to several limitations. First, the measurement intervals are not completely equal (semester vs. year) across countries. Hence, the comparability of results regarding changes in teaching behaviour across countries is limited. Furthermore, the number of measurement occasions is rather limited (2–4 occasions). When there are only a few measurement occasions there might be changes that only occurred by chance. Future ambitious longitudinal and cross-national research should try to apply equal measurement moments and more measurement occasions, preferably on a monthly basis for several years, if possible.

Second, not all countries included provided a minimum of three measurement data due to challenging circumstances. Hence, the estimation of change in some countries (Pakistan, South Africa) is limited to the linear trend only, which may not represent the true pattern of teaching behaviour change in practice.

Third, although the initial random sampling design was planned and typical lessons were observed, it is naturally difficult to avoid selection bias of teacher participants and which lessons to be observed. We caution against broader generalizations of the findings until replications of the current study are available.

Fourth, the sample and teacher characteristics across the five countries are not entirely similar. For example, Dutch samples were dominated by a high proportion of inexperienced teachers. In contrast, higher proportion of experienced teachers was more visible in the other four countries. These sample characteristics may influence the results and, thus, the current results should be interpreted with cautions until further replication studies with more representative samples are available.

Regardless of the above obvious limitations, the present study is among the first to document differences and changes in teaching behaviour using a uniform observation measure across national contexts. Despite its importance to contribute to the universal knowledge base of teaching practices, cross-national and longitudinal studies when combined are highly challenging. The present study proves that it is possible to successfully carry out this kind of ambitious research. However, adequate resources, great commitment, and dedication from multiple stakeholders are

needed, which is highly difficult to realize in typical educational research. Still, conclusions derived from such large and ambitious studies remain tentative due to the limitations mentioned. Nevertheless, the current study can pave the way toward understanding the emic and etic aspects of teaching practices that should be further investigated in the future.

Acknowledgement We are indebted to all partners who have greatly contributed to this work. We would like to specially thank Wim van de Grift (professor emeritus), who provided continuous support during the inception of the current study. We are also indebted to Anna Verkade, Geke Schuurman, and Carla Griep for their valuable contributions as international trainers. The biggest thanks goes to all teachers, schools, and observers participating in this large-scale study in the five countries. This work was supported by the Dutch scientific funding agency (NRO) under Grant number 405-15-732; the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea under Grant number NRF-2017S1A5A2A03067650, and the Directorate General of Higher Education of Indonesia under Grant number 04/SP2H/DRPM/LPPM-UNJ/III/2019.

Appendix A

1. Netherlands

Table A.1.1 MLGCM results for learning climate

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
<i>Fixed effect</i>								
Intercept	3.434***	0.016	3.064***	0.038	3.068***	0.039	3.071***	0.041
Time			0.202***	0.033	0.197***	0.033	0.195***	0.033
Time ²			-0.012*	0.006	-0.011+	0.006	-0.010+	0.006
<i>Random effects</i>								
<i>Level 3 variance (school)</i>								
Intercept	0.041	0.006	0.035	0.005	0.034	0.005	0.082	0.016
Intercept × Time							-0.018	0.004
Time							0.007	0.001
<i>Level 2 variance (Teacher)</i>								
Intercept	0.061	0.006	0.068	0.005	0.212	0.022	0.170	0.022
Intercept × Time					-0.037	0.006	-0.021	0.006
Time					0.005	0.002	0.000	0.002
<i>Level 1 variance (Time)</i>								
Residual	0.182	0.005	0.154	0.004	0.144	0.005	0.142	0.005
Deviance	5430.464		4985.748		4879.311		4847.398	

Note. + $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.1.2 MLGCM results for classroom management

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	3.280***	0.016	2.853***	0.042	2.855***	0.042	2.851***	0.044
Time			0.258***	0.037	0.255***	0.036	0.256***	0.036
Time ²			-0.023***	0.007	-0.022**	0.007	-0.022**	0.007
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.038	0.006	0.030	0.005	0.028	0.005	0.067	0.016
Intercept × Time							-0.013	0.004
Time							0.004	0.001
<i>Level 2 variance (Teacher)</i>								
Intercept	0.076	0.007	0.079	0.006	0.228	0.026	0.198	0.027
Intercept × Time					-0.044	0.008	-0.034	0.008
Time					0.011	0.003	0.007	0.003
<i>Level 1 variance (Time)</i>								
Residual	0.219	0.006	0.192	0.005	0.174	0.006	0.174	0.006
Deviance	6090.503		5705.536		5658.212		5648.860	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.1.3 MLGCM results for clarity of instruction

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	3.159***	0.015	2.740***	0.040	2.739***	0.041	2.741***	0.041
Time			0.235***	0.036	0.233***	0.035	0.233***	0.035
Time ²			-0.016*	0.007	-0.015*	0.007	-0.015*	0.007
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.032	0.005	0.027	0.005	0.028	0.005	0.030	0.012
Intercept × Time							-0.004	0.003
Time							0.003	0.001
<i>Level 2 variance (Teacher)</i>								
Intercept	0.076	0.007	0.081	0.006	0.206	0.025	0.197	0.025
Intercept × Time					-0.039	0.008	-0.033	0.008
Time					0.010	0.003	0.008	0.003
<i>Level 1 variance (Time)</i>								
Residual	0.213	0.006	0.181	0.005	0.165	0.006	0.164	0.006
Deviance	6000.573		5543.442		5510.323		5498.992	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.1.4 MLGCM results for activating teaching

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.735***	0.019	2.213***	0.045	2.213***	0.045	2.214***	0.046
Time			0.303***	0.039	0.303***	0.038	0.302***	0.039
Time ²			-0.023**	0.008	-0.023***	0.007	-0.023***	0.007
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.067	0.009	0.060	0.008	0.061	0.008	0.069	0.017
Intercept × Time							-0.008	0.005
Time							0.005	0.001
<i>Level 2 variance (Teacher)</i>								
Intercept	0.073	0.008	0.082	0.007	0.196	0.028	0.176	0.028
Intercept × Time					-0.038	0.009	-0.028	0.009
Time					0.013	0.003	0.008	0.003
<i>Level 1 variance (Time)</i>								
Residual	0.264	0.007	0.216	0.006	0.197	0.007	0.196	0.007
Deviance	6725.287		6178.167		6159.883		6142.536	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.1.5 MLGCM results for differentiated instruction

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	1.974***	0.021	1.553***	0.054	1.552***	0.052	1.554***	0.053
Time			0.263***	0.048	0.263***	0.046	0.258***	0.047
Time ²			-0.026**	0.009	-0.026**	0.009	-0.024**	0.009
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.075	0.011	0.067	0.010	0.066	0.010	0.094	0.022
Intercept × Time							-0.021	0.008
Time							0.014	0.003
<i>Level 2 variance (Teacher)</i>								
Intercept	0.096	0.010	0.103	0.010	0.175	0.037	0.126	0.036
Intercept × Time					-0.043	0.013	-0.017	0.013
Time					0.026	0.005	0.012	0.005
<i>Level 1 variance (Time)</i>								
Residual	0.344	0.010	0.318	0.009	0.281	0.010	0.281	0.010
Deviance	7696.881		7483.695		7442.237		7399.618	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.1.6 MLGCM results for teaching learning strategy

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.156***	0.025	1.738***	0.055	1.740***	0.054	1.740***	0.054
Time			0.208***	0.047	0.207***	0.046	0.201***	0.046
Time ²			-0.005	0.009	-0.006	0.009	-0.005	0.009
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.134	0.016	0.127	0.015	0.124	0.015	0.121	0.025
Intercept × Time							-0.013	0.007
Time							0.013	0.003
<i>Level 2 variance (Teacher)</i>								
Intercept	0.095	0.010	0.107	0.010	0.165	0.036	0.123	0.036
Intercept × Time					-0.033	0.013	-0.009	0.012
Time					0.020	0.005	0.007	0.005
<i>Level 1 variance (Time)</i>								
Residual	0.353	0.010	0.307	0.009	0.279	0.010	0.278	0.010
Deviance	7877.823		7521.368		7495.264		7449.895	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

2. Indonesia

Table A.2.1 MLGCM results for learning climate

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.975***	0.057	2.897***	0.133	2.897***	0.127	2.962***	0.148
Time			0.085	0.151	0.088	0.142	0.020	0.145
Time ²			-0.017	0.040	-0.019	0.038	-0.013	0.037
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.069	0.024	0.071	0.024	0.069	0.023	0.220	0.076
Intercept × Time								
Time								
<i>Level 2 variance (Teacher)</i>								
Intercept	0.085	0.016	0.084	0.014	0.249	0.075	0.154	0.068
Intercept × Time					-0.081	0.036	-0.024	0.031
Time					0.040	0.019	0.007	0.016
<i>Level 1 variance (Time)</i>								
Residual	0.177	0.015	0.177	0.014	0.146	0.017	0.151	0.017
Deviance	1098.150		1097.419		1093.669		1064.702	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.2.2 MLGCM results for classroom management

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.777***	0.088	2.895***	0.166	2.901***	0.150	2.946***	0.184
Time			-0.173	0.177	-0.180	0.152	-0.214	0.157
Time ²			0.053	0.046	0.054	0.041	0.043	0.041
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.187	0.057	0.188	0.058	0.188	0.057	0.481	0.155
Intercept × Time							-0.146	0.053
Time							0.046	0.019
<i>Level 2 variance (Teacher)</i>								
Intercept	0.102	0.021	0.104	0.021	0.511	0.101	0.392	0.092
Intercept × Time					-0.213	0.052	-0.137	0.046
Time					0.121	0.029	0.073	0.025
<i>Level 1 variance (Time)</i>								
Residual	0.243	0.019	0.241	0.019	0.146	0.019	0.153	0.019
Deviance	1320.873		1318.884		1298.878		1270.320	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$, # Time effect is not significant, none of the background variables effect is significant, thus the effect of time and background variables are nor modelled

Table A.2.3 MLGCM results for clarity of instruction

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.647***	0.081	2.961***	0.144	2.981***	0.134	3.029***	0.168
Time			-0.406***	0.108	-0.434**	0.134	-0.483***	0.140
Time ²					0.117**	0.036	0.115**	0.036
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.156	0.048	0.152	0.047	0.158	0.048	0.432	0.135
Intercept × Time							-0.139	0.049
Time							0.049	0.019
<i>Level 2 variance (Teacher)</i>								
Intercept	0.105	0.017	0.110	0.017	0.325	0.075	0.177	0.061
Intercept × Time					-0.126	0.039	-0.029	0.029
Time					0.081	0.022	0.022	0.016
<i>Level 1 variance (Time)</i>								
Residual	0.178	0.014	0.172	0.014	0.114	0.015	0.116	0.014
Deviance	1160.133		1152.890		1131.114		1082.834	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.2.4 MLGCM results for activating teaching

	Model 0		Model 1#		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.370***	0.070	2.297***	0.079	2.301***	0.080	2.305***	0.109
Time			0.052*	0.025	0.049 ⁺	0.027	0.044	0.042
Time ²								
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.115	0.036	0.116	0.036	0.119	0.037	0.255	0.086
Intercept × Time							-0.067	0.030
Time							0.024	0.012
<i>Level 2 variance (Teacher)</i>								
Intercept	0.077	0.015	0.077	0.015	0.177	0.071	0.087	0.014
Intercept × Time					-0.058	0.035	0.000	0.000
Time					0.037	0.019	0.000	0.000
<i>Level 1 variance (Time)</i>								
Residual	0.171	0.014	0.170	0.013	0.142	0.017	0.148	0.012
Deviance	1076.493		1072.254		1066.353		1038.913	

Note. ⁺ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$, # Time linear effect is significant, but the quadratic time effect is not

Table A.2.5 MLGCM results for differentiated instruction

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	1.763***	0.085	2.036***	0.162	2.060***	0.153	2.091***	0.175
Time			-0.329 ⁺	0.173	-0.361*	0.160	-0.408*	0.167
Time ²			0.080 ⁺	0.045	0.090*	0.043	0.105*	0.043
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.176	0.052	0.175	0.052	0.176	0.053	0.352	0.119
Intercept × Time							-0.094	0.044
Time							0.038	0.019
<i>Level 2 variance (Teacher)</i>								
Intercept	0.025	0.015	0.027	0.015	0.257	0.093	0.168	0.084
Intercept × Time					-0.123	0.047	-0.063	0.042
Time					0.074	0.026	0.035	0.022
<i>Level 1 variance (Time)</i>								
Residual	0.245	0.019	0.242	0.019	0.183	0.022	0.181	0.021
Deviance	1176.748		1172.645		1163.488		1144.299	

Note. ⁺ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.2.6 MLGCM results for teaching learning strategy

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.104***	0.083	2.304***	0.155	2.321***	0.150	2.378***	0.173
Time			-0.301 ⁺	0.165	-0.325*	0.157	-0.389*	0.163
Time ²			0.094*	0.043	0.102*	0.042	0.116**	0.042
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.164	0.050	0.163	0.050	0.171	0.052	0.352	0.117
Intercept × Time							-0.090	0.040
Time							0.033	0.016
<i>Level 2 variance (Teacher)</i>								
Intercept	0.076	0.018	0.081	0.018	0.201	0.087	0.092	0.078
Intercept × Time					-0.075	0.044	-0.004	0.037
Time					0.053	0.024	0.007	0.020
<i>Level 1 variance (Time)</i>								
Residual	0.218	0.017	0.212	0.017	0.172	0.021	0.177	0.021
Deviance	1214.518		1206.533		1197.156		1169.698	

Note. ⁺ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

3. South Africa

Table A.3.1 MLGCM results for learning climate

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	3.059***	0.056	2.697***	0.094	2.698***	0.098	2.714***	0.144
Time			0.243***	0.050	0.241***	0.051	0.229**	0.078
Time ²								
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.086	0.026	0.089	0.027	0.083	0.025	0.512	0.176
Intercept × Time							-0.242	0.092
Time							0.134	0.051
<i>Level 2 variance (Teacher)</i>								
Intercept	0.000	0.000	0.000	0.000	2.266	0.025	1.907	0.163
Intercept × Time					-1.297	0.107	-1.079	0.095
Time					0.808	0.065	0.675	0.058
<i>Level 1 variance (Time)</i>								
Residual	0.409	0.024	0.393	0.023	0.000	0.000	0.000	0.000
Deviance	1246.418		1224.127		1212.371		1190.616	

Note. ⁺ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.3.2 MLGCM results for classroom management

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	3.019***	0.045	2.694***	0.089	2.694***	0.093	2.714***	0.124
Time			0.218***	0.051	0.217***	0.051	0.203**	0.072
Time ²								
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.046	0.017	0.048	0.017	0.046	0.017	0.301	0.130
Intercept × Time							-0.159	0.073
Time							0.099	0.044
<i>Level 2 variance (Teacher)</i>								
Intercept	0.000	0.025	0.012	0.025	2.300	0.187	2.055	0.176
Intercept × Time					-1.304	0.108	-1.151	0.101
Time					0.812	0.066	0.716	0.061
<i>Level 1 variance (Time)</i>								
Residual	0.431	0.035	0.407	0.033	0.000	0.000	0.000	0.000
Deviance	1261.890		1244.596		1233.530		1222.120	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.3.3 MLGCM results for clarity of instruction

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.997***	0.043	2.676***	0.083	2.677***	0.091	2.689***	0.119
Time			0.215***	0.048	0.214***	0.050	0.205**	0.074
Time ²								
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.044	0.015	0.045	0.015	0.050	0.016	0.278	0.119
Intercept × Time							-0.162	0.072
Time							0.114	0.046
<i>Level 2 variance (Teacher)</i>								
Intercept	0.000	0.000	0.000	0.000	2.157	0.175	1.887	0.161
Intercept × Time					-1.258	0.103	-1.084	0.095
Time					0.787	0.063	0.675	0.058
<i>Level 1 variance (Time)</i>								
Residual	0.365	0.021	0.353	0.020	0.000	0.000	0.000	0.000
Deviance	1162.844		1143.172		1125.882		1108.686	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.3.4 MLGCM results for activating teaching

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.883***	0.044	2.542***	0.084	2.541***	0.091	2.547***	0.111
Time			0.229***	0.047	0.229***	0.047	0.225***	0.066
Time ²								
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.046	0.016	0.048	0.016	0.051	0.016	0.214	0.105
Intercept × Time							-0.114	0.059
Time							0.079	0.036
<i>Level 2 variance (Teacher)</i>								
Intercept	0.000	0.000	0.001	0.020	2.111	0.171	1.909	0.163
Intercept × Time					-1.165	0.096	-1.040	0.091
Time					0.693	0.056	0.615	0.053
<i>Level 1 variance (Time)</i>								
Residual	0.364	0.021	0.348	0.028	0.000	0.000	0.000	0.000
Deviance	1161.113		1138.799		1103.427		1092.157	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.3.5 MLGCM results for differentiated instruction

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.732***	0.053	2.233***	0.104	2.234***	0.111	2.238***	0.156
Time			0.335***	0.060	0.334***	0.062	0.330***	0.092
Time ²								
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.066	0.024	0.070	0.024	0.070	0.023	0.534	0.205
Intercept × Time							-0.291	0.118
Time							0.183	0.072
<i>Level 2 variance (Teacher)</i>								
Intercept	0.000	0.000	0.000	0.000	3.199	0.259	2.735	0.234
Intercept × Time					-1.865	0.154	-1.572	0.138
Time					1.179	0.095	0.994	0.085
<i>Level 1 variance (Time)</i>								
Residual	0.583	0.034	0.553	0.032	0.000	0.000	0.000	0.000
Deviance	1447.529		1417.473		1408.139		1388.790	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.3.6 MLGCM results for teaching learning strategy

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.872***	0.045	2.563***	0.089	2.563***	0.097	2.577***	0.125
Time			0.208***	0.052	0.207***	0.053	0.196***	0.083
Time ²								
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.046	0.017	0.048	0.017	0.055	0.018	0.307	0.132
Intercept × Time							-0.201	0.084
Time							0.154	0.057
<i>Level 2 variance (Teacher)</i>								
Intercept	0.000	0.000	0.000	0.000	2.414	0.196	2.079	0.178
Intercept × Time					-1.396	0.115	-1.173	0.103
Time					0.873	0.071	0.725	0.062
<i>Level 1 variance (Time)</i>								
Residual	0.454	0.024	0.413	0.024	0.000	0.000	0.000	0.000
Deviance	1251.788		1236.101		1233.916		1199.217	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

4. Mongolia

Table A.4.1 MLGCM results for learning climate

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	3.114***	0.043	2.637***	0.122	2.639***	0.103	2.645***	0.116
Time			0.662***	0.130	0.662***	0.108	0.658***	0.112
Time ²			-0.181***	0.032	-0.181***	0.026	-0.181***	0.026
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.071	0.018	0.071	0.018	0.067	0.017	0.215	0.057
Intercept × Time							-0.087	0.026
Time							0.051	0.014
<i>Level 2 variance (Teacher)</i>								
Intercept	0.039	0.011	0.044	0.011	0.134	0.050	0.010	0.044
Intercept × Time					-0.100	0.025	-0.023	0.021
Time					0.086	0.014	0.038	0.011
<i>Level 1 variance (Time)</i>								
Residual	0.277	0.014	0.262	0.013	0.175	0.012	0.175	0.012
Deviance	1959.215		1917.182		1805.979		1750.051	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.4.2 MLGCM results for classroom management

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.994***	0.044	2.662***	0.119	2.663***	0.101	2.667***	0.104
Time			0.463***	0.126	0.463***	0.105	0.460***	0.107
Time ²			-0.127***	0.031	-0.127***	0.025	-0.127***	0.025
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.072	0.019	0.072	0.019	0.061	0.017	0.089	0.034
Intercept × Time							-0.029	0.015
Time							0.025	0.009
<i>Level 2 variance (Teacher)</i>								
Intercept	0.065	0.012	0.068	0.012	0.206	0.052	0.147	0.050
Intercept × Time					-0.107	0.024	-0.069	0.023
Time					0.080	0.013	0.056	0.012
<i>Level 1 variance (Time)</i>								
Residual	0.253	0.013	0.246	0.012	0.165	0.012	0.165	0.012
Deviance	1937.000		1914.433		1834.353		1811.384	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.4.3 MLGCM results for clarity of instruction

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.941***	0.041	2.569***	0.112	2.570***	0.098	2.572***	0.103
Time			0.503***	0.119	0.503***	0.103	0.502***	0.105
Time ²			-0.136***	0.029	-0.136***	0.025	-0.136***	0.025
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.061	0.017	0.061	0.017	0.056	0.015	0.104	0.034
Intercept × Time							-0.031	0.014
Time							0.020	0.007
<i>Level 2 variance (Teacher)</i>								
Intercept	0.063	0.011	0.066	0.011	0.110	0.046	0.062	0.044
Intercept × Time					-0.061	0.021	-0.031	0.020
Time					0.055	0.011	0.036	0.010
<i>Level 1 variance (Time)</i>								
Residual	0.226	0.011	0.218	0.011	0.162	0.011	0.162	0.011
Deviance	1817.924		1791.597		1724.514		1708.220	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.4.4 MLGCM results for activating teaching

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.756***	0.042	2.33***3	0.109	2.337***	0.095	2.349***	0.103
Time			0.528***	0.115	0.528***	0.097	0.521***	0.100
Time ²			-0.136***	0.028	-0.136***	0.023	-0.136***	0.023
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.069	0.018	0.069	0.018	0.066	0.017	0.148	0.043
Intercept × Time							-0.048	0.017
Time							0.028	0.009
<i>Level 2 variance (Teacher)</i>								
Intercept	0.053	0.010	0.055	0.010	0.156	0.044	0.084	0.040
Intercept × Time					-0.080	0.020	-0.037	0.018
Time					0.061	0.010	0.035	0.009
<i>Level 1 variance (Time)</i>								
Residual	0.210	0.010	0.203	0.010	0.142	0.010	0.142	0.010
Deviance	1727.654		1704.487		1637.958		1609.461	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.4.5 MLGCM results for differentiated instruction

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.360***	0.048	1.584***	0.111	1.585***	0.099	1.596***	0.113
Time			0.906***	0.115	0.906***	0.097	0.900***	0.100
Time ²			-0.222***	0.028	-0.222***	0.023	-0.222***	0.023
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.094	0.023	0.094	0.023	0.096	0.023	0.250	0.066
Intercept × Time							-0.071	0.023
Time							0.032	0.009
<i>Level 2 variance (Teacher)</i>								
Intercept	0.057	0.010	0.062	0.010	0.260	0.050	0.155	0.045
Intercept × Time					-0.105	0.022	-0.051	0.019
Time					0.060	0.010	0.033	0.009
<i>Level 1 variance (Time)</i>								
Residual	0.220	0.011	0.203	0.010	0.142	0.010	0.142	0.010
Deviance	1794.227		1734.951		1695.444		1662.245	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.4.6 MLGCM results for teaching learning strategy

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.726***	0.046	2.334***	0.110	2.334***	0.100	2.328***	0.105
Time			0.459***	0.115	0.459***	0.101	0.462***	0.102
Time ²			-0.112***	0.028	-0.112***	0.024	-0.112***	0.024
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.087	0.021	0.087	0.021	0.084	0.020	0.136	0.041
Intercept × Time							-0.028	0.014
Time							0.015	0.006
<i>Level 2 variance (Teacher)</i>								
Intercept	0.047	0.009	0.048	0.009	0.137	0.045	0.089	0.044
Intercept × Time					-0.067	0.020	-0.040	0.019
Time					0.049	0.010	0.033	0.010
<i>Level 1 variance (Time)</i>								
Residual	0.208	0.010	0.204	0.010	0.155	0.011	0.155	0.011
Deviance	1713.237		1697.593		1654.475		1639.463	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

5. Pakistan

Table A.5.1 MLGCM results for learning climate

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.187***	0.060	1.711***	0.079	1.708***	0.085	1.683***	0.206
Time			0.317***	0.034	0.317***	0.036	0.335**	0.112
Time ²								
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.059	0.021	0.059	0.021	0.084	0.029	0.737	0.255
Intercept × Time							-0.386	0.137
Time							0.214	0.075
<i>Level 2 variance (Teacher)</i>								
Intercept	0.000	0.000	0.000	0.000	0.906	0.070	0.519	0.041
Intercept × Time					-0.606	0.048	-0.326	0.027
Time					0.448	0.034	0.245	0.019
<i>Level 1 variance (Time)</i>								
Residual	0.231	0.012	0.205	0.011	0.000	0.000	0.000	0.000
Deviance	966.027		888.329		866.253		702.977	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.5.2 MLGCM results for classroom management

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.076***	0.069	1.823***	0.088	1.822***	0.091	1.802***	0.228
Time			0.168***	0.036	0.168***	0.039	0.182	0.127
Time ²								
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.080	0.028	0.080	0.028	0.090	0.031	0.906	0.312
Intercept × Time							-0.484	0.170
Time							0.278	0.097
<i>Level 2 variance (Teacher)</i>								
Intercept	0.000	0.000	0.000	0.000	1.157	0.089	0.560	0.044
Intercept × Time					-0.738	0.058	-0.329	0.027
Time					0.512	0.039	0.231	0.018
<i>Level 1 variance (Time)</i>								
Residual	0.230	0.012	0.222	0.012	0.000	0.000	0.000	0.000
Deviance	967.752		946.605		935.925		720.437	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.5.3 MLGCM results for clarity of instruction

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.076***	0.066	1.796***	0.084	1.791***	0.090	1.773***	0.197
Time			0.186***	0.034	0.186***	0.035	0.201+	0.111
Time ²								
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.074	0.026	0.075	0.026	0.108	0.037	0.680	0.233
Intercept × Time							-0.360	0.128
Time							0.211	0.074
<i>Level 2 variance (Teacher)</i>								
Intercept	0.000	0.000	0.000	0.000	0.736	0.057	0.394	0.031
Intercept × Time					-0.520	0.041	-0.252	0.021
Time					0.414	0.031	0.204	0.016
<i>Level 1 variance (Time)</i>								
Residual	0.208	0.011	0.199	0.011	0.000	0.000	0.000	0.000
Deviance	900.386		871.649		827.061		635.306	

Note. + $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.5.4 MLGCM results for activating teaching

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.021***	0.057	1.788***	0.070	1.787***	0.069	1.768***	0.134
Time			0.154***	0.027	0.154***	0.027	0.168*	0.079
Time ²								
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.055	0.019	0.056	0.019	0.059	0.020	0.308	0.108
Intercept × Time							-0.163	0.061
Time							0.105	0.037
<i>Level 2 variance (Teacher)</i>								
Intercept	0.000	0.000	0.003	0.007	0.525	0.040	0.330	0.026
Intercept × Time					-0.337	0.026	-0.196	0.016
Time					0.246	0.019	0.144	0.011
<i>Level 1 variance (Time)</i>								
Residual	0.133	0.007	0.123	0.009	0.000	0.000	0.000	0.000
Deviance	603.521		572.310		559.450		426.826	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.5.5 MLGCM results for differentiated instruction

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	2.003***	0.064	1.744***	0.081	1.773***	0.080	1.743***	0.170
Time			0.152***	0.033	0.152***	0.033	0.172	0.105
Time ²								
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.068	0.024	0.069	0.024	0.069	0.024	0.495	0.174
Intercept × Time							-0.284	0.104
Time							0.189	0.066
<i>Level 2 variance (Teacher)</i>								
Intercept	0.000	0.000	0.000	0.000	0.879	0.068	0.501	0.039
Intercept × Time					-0.548	0.043	-0.286	0.023
Time					0.380	0.029	0.199	0.015
<i>Level 1 variance (Time)</i>								
Residual	0.193	0.010	0.187	0.010	0.000	0.000	0.000	0.000
Deviance	849.950		829.395		826.489		661.585	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A.5.6 MLGCM results for teaching learning strategy

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Fixed effect								
Intercept	1.922***	0.059	1.599***	0.074	1.598***	0.073	1.571***	0.158
Time			0.215***	0.030	0.215***	0.030	0.233*	0.099
Time ²								
Random effects								
<i>Level 3 variance (school)</i>								
Intercept	0.059	0.021	0.059	0.021	0.059	0.021	0.429	0.150
Intercept × Time							−0.252	0.091
Time							0.172	0.059
<i>Level 2 variance (Teacher)</i>								
Intercept	0.000	0.000	0.000	0.000	0.718	0.056	0.372	0.029
Intercept × Time					−0.447	0.035	−0.207	0.017
Time					0.310	0.023	0.144	0.011
<i>Level 1 variance (Time)</i>								
Residual	0.164	0.009	0.153	0.008	0.000	0.000	0.000	0.000
Deviance	744.105		695.118		692.256		492.287	

Note. * $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

References

- Ahmad, I., Rehman, K., Ali, A., Khan, I., & Khan, F. A. (2014). Critical analysis of the problems of education in Pakistan. *International Journal of Evaluation and Research in Education*, 3(2), 79–84.
- Baller, S., Dutta, S., & Lanvin, B. (Eds.). (2016). *The global information technology report 2016: Innovating in the digital economy*. World Economic Forum. http://www3.weforum.org/docs/GITR2016/WEF_GITR_Full_Report.pdf
- Brophy, J. (1986). Teacher influences on student achievement. *American Psychologist*, 41(10), 1069–1077.
- Bruns, B., De Gregorio, S., & Taut, S. (2016). Measures of effective teaching in developing countries. *Research on Improving Systems of Education (RISE) Working Paper*, 16(9), 1–37. https://riseprogramme.org/sites/default/files/publications/RISE_WP-009_Bruns_0.pdf
- Chun, S., Galindez, U., Kim, D., Adiyasuren, A., & Kim, J. (2020). A study on the applicability of ICALT for comparing teaching skills between secondary schools in Korea and Mongolia. *The SNU Journal of Education Research*, 29(1), 53–75.
- Coe, R., Aloisi, C., Higgins, S., & Elliott Major, L. (2014). *What makes great teaching? A review of the underpinning research*. The Sutton Trust.
- Danielson, C. (2013). *The framework for teaching evaluation instrument*. The Danielson Group.
- Fuller, F. (1970). *Personalized education for teachers: One application of the teachers concerns model*. R & D Center for Teacher Education.
- Givvin, K. B., Hiebert, J., Jacobs, J. K., Hollingsworth, H., & Gallimore, R. (2005). Are there national patterns of teaching? Evidence from the TIMSS 1999 video study. *Comparative Education Review*, 49(3), 311–343.

- Hambleton, R. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment, 10*, 229–244.
- Hebert, E., & Worthy, T. (2001). Does the first year of teaching have to be a bad one? A case study of success. *Teaching and Teacher Education, 17*, 897–911. [https://doi.org/10.1016/S0742-051X\(01\)00039-7](https://doi.org/10.1016/S0742-051X(01)00039-7)
- Hoy, W. K., & Woolfolk, A. E. (1990). Socialization of student teachers. *American Educational Research Journal, 27*, 279–300.
- Khan, W. (2019). Quality of teacher education in Pakistan. *The Dialogue, 10*(2). Retrieved May 13, 2021, from https://www.qurtuba.edu.pk/thedialogue/The%20Dialogue/10_2/Dialogue_April_June2015_212-219.pdf
- Kunter, M., & Baumert, J. (2006). Linking TIMSS to research on learning and instruction: A re-analysis of the German TIMSS and TIMSS video data. In S. J. Howie & T. Plomp (Eds.), *Learning mathematics and science: Lessons learned from TIMSS* (pp. 335–351). Routledge.
- Lumadi, M. W. (2008). Teachers' exodus in South African schools. *Contemporary Issues in Education Research, 1*(3), 31–40.
- Mainhard, M. T., Brekelmans, J. M. G., den Brok, P. J., & Wubbels, T. (2011). The development of the classroom social climate during the first months of the school year. *Contemporary Educational Psychology, 36*(3), 190–200. <https://doi.org/10.1016/j.cedpsych.2010.06.002>
- Malmberg, L. E., Hagger, H., Burn, K., Mutton, T., & Colls, H. (2010). Observed classroom quality during teacher education and two years of professional practice. *Journal of Educational Psychology, 102*(4), 916–932. <https://doi.org/10.1037/a0020920>
- Maulana, R. (2012). *Teacher-student relationships during the first year of secondary education: Exploration of change and link with motivational outcomes in the Netherlands and Indonesia*. University of Groningen, Faculty of Behavioral and Social Sciences.
- Maulana, R., & Helms-Lorenz, M. (2016). Observations and student perceptions of the quality of preservice teachers' teaching behaviour: Construct representation and predictive quality. *Learning Environments Research, 19*, 335–357. <https://doi.org/10.1007/s10984-016-9215-8>
- Maulana, R., Opendakker, M.-C., den Brok, P., & Bosker, R. (2011). Teacher-student interpersonal relationships in Indonesia: Profiles and importance to student motivation. *Asia Pacific Journal of Education, 31*(1), 33–49. <https://doi.org/10.1080/02188791.2011.544061>
- Maulana, R., Opendakker, M.-C., Stroet, K., & Bosker, R. (2012). Observed lesson structure during the first year of secondary education: Exploration of change and link with academic engagement. *Teaching and Teacher Education, 28*(6), 835–850. <https://doi.org/10.1016/j.tate.2012.03.005>
- Maulana, R., Opendakker, M.-C., Stroet, K., & Bosker, R. (2013). Changes in teachers' involvement versus rejection and links with academic motivation during the first year of secondary education: A multilevel growth curve analysis. *Journal of Youth Adolescence, 42*, 1348–1371. <https://doi.org/10.1007/s10964-013-9921-9>
- Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2015). A longitudinal study of induction on the acceleration of growth in teaching quality of beginning teachers through the eyes of their students. *Teaching and Teacher Education, 51*, 225–245. <https://doi.org/10.1016/j.tate.2015.07.003>
- Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2017). Validating a model of effective teaching behaviour of pre-service teachers. *Teachers and Teaching: Theory and Practice, 23*(4), 471–493. <https://doi.org/10.1080/13540602.2016.1211102>
- Maulana, R., André, S., Helms-Lorenz, M., Ko, J., Chun, S., Shahzad, A., Irnidayanti, Y., Lee, O., de Jager, T., Coetzee, T., & Fadhilah, N. (2020a). Observed teaching behaviour in secondary education across six countries: Measurement invariance and indication of cross-national variations. *School Effectiveness and School Improvement, 1–32*, 64–95. <https://doi.org/10.1080/09243453.2020.1777170>
- Maulana, R., Smale-Jacobse, A., Helms-Lorenz, M., Chun, S., & Lee, O. (2020b). Measuring differentiated instruction in The Netherlands and South Korea: Factor structure equivalence, cor-

- relates, and complexity level. *European Journal of Psychology of Education*, 35(4), 881–909. <https://doi.org/10.1007/s10212-019-00446-4>
- Mbiti, I. M. (2016). The need for accountability in education in developing countries. *The Journal of Economic Perspectives*, 30(3), 109–132.
- Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art—teacher effectiveness and professional learning. *School Effectiveness and School Improvement*, 25(2), 231–256. <https://doi.org/10.1080/09243453.20192011.4.885451>
- Muijs, D., Reynolds, D., Sammons, P., Kyriakides, L., Creemers, B. P. M., & Teddlie, C. (2018). Assessing individual lessons using a generic teacher observation instrument: How useful is the International System For Teacher Observation And Feedback (ISTOF)? *ZDM Mathematics Education*, 50(3), 395–406. <https://doi.org/10.1007/s11858-018-0921-9>
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 international results in reading*. Boston College, TIMSS & PIRLS International Study Center. <http://timssandpirls.bc.edu/pirls2016/international-results/>
- OECD. (2016a). *Netherlands 2016: Foundations for the future. Reviews of policies for national education*. OECD. <https://doi.org/10.1787/9789264257658-en>
- OECD. (2016b). *Country Note – Results from PISA 2015: Indonesia*. Retrieved from: <https://www.oecd.org/pisa/PISA-2015-Indonesia.pdf>
- OECD. (2018). *PISA 2015 Results in focus*. Retrieved from: <http://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf>
- OECD & ADB. (2015). *Education in Indonesia: Rising to the challenge* (Technical report). OECD.
- Opendakker, M.-C., Maulana, R., & den Brok, P. (2012). Teacher-student interpersonal relationships and academic motivation within one school year: Developmental changes and linkage. *School Effectiveness and School Improvement*, 23, 95–119. <https://doi.org/10.1080/09243453.2011.619198>
- Pianta, R. C., Belsky, J., Vandergrift, N., Hours, R., & Morrison, F. (2008). Classroom effects on children's achievement trajectories in elementary school. *American Educational Research Journal*, 45, 365–397. <https://doi.org/10.3102/0002831207308230>
- Rasbash, J., Charlton, C., Jones, K., & Pillinger, R. (2014). *Manual supplement for MLwiN Version 2.3.1*. <http://www.bristol.ac.uk/cmm/media/software/mlwin/downloads/manuals/2-33/supplement-print.pdf>
- Reynolds, D., Creemers, B., Stringfield, S., Teddlie, C., & Schaffer, G. (Eds.). (2002). *World class schools: International perspectives in school effectiveness*. Routledge Falmer.
- Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher-student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research*, 81, 493–529.
- Sandilos, L. E., Sims, W. A., Norwalk, K. E., & Reddy, L. A. (2019). Converging on quality: Examining multiple measures of teaching effectiveness. *Journal of School Psychology*, 74, 10–28.
- Shuell, T. J. (1996). Teaching and learning in a classroom context. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 726–764). Macmillan.
- Smale-Jacobse, A., Meijer, A., Helms-Lorenz, M., & Maulana, R. (2019). Differentiated instruction in secondary education: A systematic review and meta-analysis. *Frontiers in Psychology*, 10, 2366. <https://doi.org/10.3389/fpsyg.2019.02366>
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel-modeling* (2nd ed.). Sage Publishers.
- Stallings, J. (1977). *Learning to look: A handbook on classroom observation and teaching models*. Wadsworth Publishing.
- Stallings, J., & Mohlman, G. (1988). Classroom observation techniques. In J. Keeves (Ed.), *Educational research, methodology, and measurement: An international handbook* (pp. 469–474). Pergamon Press.
- Steiner-Khamsi, G., & Gerelmaa, A. (2008). Quality and equity in the Mongolian education sector. *Prospects*, 38, 409–414.

- Stigler, J. W., Gonzales, P. A., Kawanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS video-tape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States*. U.S. Department of Education, National Center for Education Statistics.
- Stroet, K., Opdenakker, M. C., & Minnaert, A. (2015). What motivates early adolescents for school? A longitudinal analysis of associations between observed teaching and motivation. *Contemporary Educational Psychology*, 42, 129–140. <https://doi.org/10.1016/j.cedpsych.2015.06.002>
- United Nations Educational, Scientific and Cultural Organization. (2006). *Situation analysis of teacher education: Towards a strategic framework for teacher education and professional development*.
- United Nations Educational, Scientific and Cultural Organization. (2015). *Regional overview: South and West Asia*.
- United Nations Educational, Scientific and Cultural Organization. (2020). *Mongolia Education Policy review: Towards a lifelong learning system*. <https://unesdoc.unesco.org/ark:/48223/pf00000373687>
- van de Grift, W. J. (2014). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement*, 25(3), 295–311.
- van de Grift, W. J., Helms-Lorenz, M., & Maulana, R. (2014). Teaching skills of student teachers: Calibration of an evaluation instrument and its value in predicting student academic engagement. *Studies in Educational Evaluation*, 43, 150–159. <https://doi.org/10.1016/j.stueduc.2014.09.003>
- van de Grift, W. J., Chun, S., Maulana, R., Lee, O., & Helms-Lorenz, M. (2017). Measuring teaching quality and student engagement in South Korea and The Netherlands. *School Effectiveness and School Improvement*, 28(3), 337–349. <https://doi.org/10.1080/09243453.2016.1263215>
- van der Berg, S., & Hofmeyr, H. (2017). *South African education: Background note for World Bank's South African country diagnostic, research on socio-economic policy*. World Bank. Retrieved from <http://documents1.worldbank.org/curated/en/339291529320964248/pdf/127304-Education-in-South-Africa.pdf>
- Virtanen, T. E., Lerkkanen, M.-J., Poikkeus, A.-M., & Kuorelahti, M. (2015). The relationship between classroom quality and students' engagement in secondary school. *Educational Psychology*, 35(8), 963–983. <https://doi.org/10.1080/01443410.2013.822961>
- Woolfolk Hoy, A., & Burke Spero, R. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and Teacher Education*, 21, 343–356. <https://doi.org/10.1016/j.tate.2005.01.007>
- Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (1997). *Program evaluation: Alternative approaches and practical guidelines*. Longman.

Ridwan Maulana is an associate professor at the Department of Teacher Education, University of Groningen, the Netherlands. His major research interests include teaching and teacher education, factors influencing effective teaching, methods associated with the measurement of teaching, longitudinal research, cross-country comparisons, effects of teaching behaviour on students' motivation and engagement, and teacher professional development. He has been involved in various teacher professional development projects including the Dutch induction programme and school–university-based partnership. He is currently a project leader of an international project on teaching quality involving countries from Europe, Asia, Africa, Australia, and America. He is a European Editor of Learning Environments Research journal, a SIG leader of Learning Environments of American Educational Research Association, and chair of the Ethics Commission of the Teacher Education.

Amanda Maraschin Bruscato holds a PhD in Language Sciences from the University of Algarve, Portugal, and currently works as an English teacher at the Montessori Lyceum Groningen, the Netherlands. Her research interests include teaching effectiveness, second language learning, and e-learning.

Michelle Helms-Lorenz is an Associate Professor at the Department of Teacher Education, University of Groningen, The Netherlands. Her research interest covers the cultural specificity versus universality (of behaviour and psychological processes). This interest was fed by the cultural diversity in South Africa, where she was born and raised. Michelle's second passion is education, the bumpy road toward development. Her research interests include teaching skills and well-being of beginning and pre-service teachers and effective interventions to promote their professional growth and retention.

Yulia Irmidayanti obtained her first degree in Biology Education and PhD in Biology. She is currently a Senior Lecturer and researcher at the Biology and Biology Education Department, Universitas Negeri Jakarta [State University of Jakarta], Indonesia. Since 2001, she has been working together with the Teacher Education Department of University of Groningen, the Netherlands, on the project about teaching quality and student academic motivation from the international perspective (ICALT3/Differentiation project, Principal investigator Indonesia). She is interested in helping teachers to improve their teaching quality and student differences in their learning needs, motivation, and learning style.

Thelma de Jager is the HOD of the Department Educational Foundation and a senior lecturer. She received several awards for woman researcher and lecturer of the year, conducted several keynote addresses at conferences and authored and edited textbooks such as: *General Subject Didactics*, *Creative Arts Education The Science to Teach and Differentiated Instruction*. She is currently the project leader of the South African team for the ICALT 3 project and the British Council on Inclusive Education, T4ALL project. She has a passion to improve teaching pedagogy and her studies could impact education policy that speaks to the implementation of differentiated learning.

Ulziisaikhan Galindev is a senior lecturer in The Department of Educational Administration, Mongolian National University of Education. He received his master and doctoral degrees in Educational administration from Chungnam National University, South Korea. His current research interests and expertise cover education finance, education policy and teacher professional development.

Amarjargal Adiyasuren is a lecturer at Mongolian National University of Education. She formerly worked in Teachers' Professional Development Institute and Curriculum Reform Unit affiliated to Ministry of Education and Science. She worked in various national research projects related to school management, curriculum, pedagogy and assessment. She has been involved in comparative study of assessment of transversal skills with the Network on Education Quality Monitoring in the Asia-Pacific in the UNESCO Asia-Pacific and the Brookings Institution of the USA. She holds bachelor and master degree in Education from the University of Tokyo.

Dr. Abid Shahzad is the Founding Director of the International Linkages at the Islamia University of Bahawalpur (IUB). He earned his PhD degree in Educational Sciences from Ghent University Belgium. Currently, he is serving as an Assistant Professor at the Department of Education. He has presented his research papers and participated in international workshops and seminars in more than twenty countries. He is also the founder of the International Conference on Teaching and Learning (ICOTAL) and International STEMS conference that are held annually at the Islamia University of Bahawalpur. He has engaged a number of renowned international universities, research institutes and educationists in the ICOTAL and STEMS conferences. He is regularly organizing international training workshops for research students at the Faculty of Education. He is actively signing MoUs with international academic and research partners.

Nurul Fadhilah is a university lecturer at the Department of Biostatistic and Population, University of Indonesia. She has been actively involved in the international project called ICALT3/ Differentiation as an expert observer and as coinvestigator for Indonesia. She is currently involved in a research project involving public health big data analysis. She has been involved in professional teacher development for high school teachers in DKI Jakarta. She is experienced in designing and facilitating teacher professional development training, developing syllabus, task designing, developing differentiated instructions, especially in Cambridge IGCSE and A level Biology subject.

Prof. Seyeoung Chun is a professor (emeritus) at the Department of Education of the Chungnam National University of South Korea since 1997. He received his education and PhD from Seoul National University and has been actively doing research in educational policy and has had several key positions, such as Secretary of Education to the President and CEO of the Korean Educational research and Information Service (KERIS). Professor Chun is the founder and current president of Smart Education Society with more than 3,000 members.

Okhwa Lee is a professor (emeritus) at the Department of Education, Chungbuk National University, South Korea, and CEO of SmartSchool (Ltd). Okhwa Lee is a specialist in educational technology and a practitioner in pre-service teacher education. She is a pioneer of software education, e-learning, and smart education in Korea. She was a member of the Presidential Educational Reform Committee and the Presidential e-Government. She has collaborated in the European Erasmus mobility programme, in research with Finland and The Netherlands, and in the Korean government ODA (Official Development Assistant) programme for Nigeria, Vietnam, and Ethiopia.

Mattheus (Thys) Coetzee is the Head of the Department Multimedia Design and Development, Tshwane University of Technology, South Africa. His research interests include teacher education, retention and success rates in engineering studies as well as statistics of student success. His fields of expertise include academic leadership in higher education, implementation of online teaching and learning technologies and the development of multimedia for teaching and learning. He was a project leader on an international project to implement eLearning in South African higher education and currently he leads an Academic Leadership project with two Dutch universities and Southern African institutions.

Peter Moorer is a former staff member of the Department of Teacher Education, University of Groningen. He worked for projects researching induction tracts for starting teachers (BSL Project) and ICALT3. He built a complex data system for the data collection on 3000 starting teachers, different questionnaire and 50 observers. His research interest is in statistics and theoretical human sciences (psychology, medicine, sociology and economics).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 18

Teacher-Student Interactions: Theory, Measurement, and Evidence for Universal Properties That Support Students' Learning Across Countries and Cultures



Tara Hofkens, Robert C. Pianta, and Bridget Hamre

Abstract Across the globe, strategies and investments to strengthen teacher effectiveness are increasingly a core component of countries' efforts to improve educational outcomes for their citizens and, for many, to elevate standards of living. In this chapter, we present evidence demonstrating the role of teacher-student interactions in teachers' ability to positively influence student development and learning across countries and cultures. We conceptualize teacher-student interactions as proximal processes that drive students' engagement and learning. Evidence clearly demonstrates that interactions can be assessed through observation and improved through professional development interventions. Drawing on our experience and data available on tens of thousands of classroom observations across different countries and cultures, we present a framework that describes core features of effective teacher-student interactions that appear in common across these highly varied settings and cultural contexts. We review research that evaluates this framework in different contexts to examine the effects of interaction quality on student outcomes across the globe. We discuss the cross-cultural applicability of the framework and outline suggestions for education policy and practice and future directions for research.

Keywords Teacher-student interactions · Classroom quality · Teaching through interactions

T. Hofkens (✉) · R. C. Pianta · B. Hamre
School of Education and Human Development, University of Virginia,
Charlottesville, VA, USA
e-mail: th7ub@virginia.edu

© The Author(s) 2023
R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_18

399

1 Introduction

In nearly all theories of education and its impacts, the quality of students' experiences in the classroom (or childcare) setting is often described as a critical, if not necessary, factor in determining the value of education. In numerous studies of educational "inputs" intended to promote student learning (e.g., funding, class size, teacher qualifications, curriculum), over and above students' prior performance and family background (Nye et al., 2004; Reardon et al., 2013). Such large-scale efforts reinforce the idea that the quality of what takes place in classrooms may be the essential ingredient for fostering student success (Heckman, 2000).

In fact, our and others' research (see Morrison & Connor, 2002; Pianta et al., 2007; Sanders & Rivers, 1996) has generated a set of generally-accepted findings and observations about teachers and teaching, albeit largely based on data collected from U.S. and Western society classrooms: (1) teachers are the most potent asset the education system provides to foster student learning and development (Sabol et al., 2013); (2) qualities of teacher-student interactions that foster student engagement and effort, knowledge and thinking, problem-solving and communication skills, and positive relationships with others are the source of these teacher effects (Pianta & Allen, 2008); (3) these qualities of teachers' interactions can be observed and measured, and predict student students' development across a range of indicators (Allen et al., 2011; Pianta et al., 2008); (4) effective teaching can be learned, trained, and improved; and (5) ensuring effective teaching at scale requires workforce development systems that integrate description, measurement, improvement, and implementation support (Pianta et al., 2020).

These conclusions are not just the result of scientific studies conducted by academics. In experience accumulated from working to assess and improve teacher-student interactions at large scale over the past decade (5000 coaches, 17,000 observers trained to agreement on CLASS, 50 countries and all 50 U.S. states), practitioners and policymakers alike describe the unique value created when teachers and their interactions with students are elevated as a developmental and educational resource.

In the present chapter we draw from cross-cultural observations of classrooms using the CLASS to evaluate the extent to which there may be patterns and features of teacher-student interaction that have common value for student learning and development. We draw from countries has varied as Sweden (Castro et al., 2017), Ecuador (Carneiro et al., 2019), and China (Hu et al., 2016) in an effort to capture the relevance of teacher-student interaction across cultures. Also, emerging evidence from international work (e.g., Carneiro et al., 2019) that supports theory of the universality of adult-child interactions for promoting development (<https://www.oecd-ilibrary.org/sites/617837e6-en/index.html?itemId=/content/component/617837e6-en>).

2 Theoretical Framework

To begin, we frame some of the terms used in the discussion. Clearly, the term “international” could have varied meanings (Maulana et al., 2021). For example, for studies pertaining only to CLASS and not to other observational instruments, “international” applications include countries as wide-ranging as Finland, Israel, Kazakhstan, Australia, and Ecuador. Important efforts to understand those sources of variance have revealed not only the complexities of assessing teacher effectiveness cross-nationally, but that there is also evidence of commonalities (e.g., Maulana et al., 2021). Rather, our aim is to advance theoretical perspectives on education and human development that posit the importance of relationships between teachers and students (Ryan & Deci, 2000). Furthermore, we recognize the widely varying nature of “teachers” and “classrooms” in the countries and cultures we include in this analysis. The data available on classroom interactions, particularly when CLASS has been the assessment, skews toward younger ages and U.S. settings, although not exclusively. Accordingly, we will make an effort to present a balanced and well-informed picture.

2.1 *Defining Effective Teaching*

In a sense, every “measure” of educational quality and opportunity is actually a test of a theory; in considering effective teaching as reflective of educational opportunity, each measure of effective teaching is a set of hypotheses about the process of teaching and learning. Each measure also reflects a set of hypotheses about how to best gather information on the construct of interest, and when a measure is used in the field the resulting data provide a form of confirmation or disconfirmation of the underlying hypotheses and theory. CLASS has been anchored in the science and theory of human development in which proximal processes between individuals are posited to account for students’ growth in broad areas of development, including cognition, achievement, social relationships, self-regulation, motivation, and identity (Bronfenbrenner & Morris, 1998). This conceptual basis drew heavily from theories of human attachment and parent-child relationships (and associated measures) to conceptualize teacher-student interactions and relationships and embarked on studies examining how best to apply this work in classrooms (Pianta, 1999).

There is little question that teachers and their classroom interactions with students matter for student achievement (Carneiro et al., 2019; Goe, 2007; Hu et al., 2016; Kane & Staiger, 2012; Loeb et al., 2012), motivation (Patrick et al., 2001; Ruzek et al., 2016; Wang & Holcombe, 2010), and a range of behavioral and social outcomes (Hoang et al., 2018; Pakarinen et al., 2011; Wang et al., 2020). Efforts to describe effective teaching have been reported in a large number of small-sample studies, and in narrative descriptions that lack evidence of validity or tools for data collection (Lemov, 2010). It has also been challenging to define and measure the

aspects of teacher behavior unique to teaching a certain content area (Hill, 2010; Grossman et al., 2014; van Hover et al., 2012) or grade level (Pianta, 2016). Measures of the same construct also vary with respect to differential suitability for data collection methods such as observation or informant report (Raudenbush & Jean, 2014; Ruzek et al., 2016; Kane et al., 2014).

When studies have included different approaches to assessing teacher-student interactions, such as evaluating multiple observation tools (Kane & Staiger, 2012; Staiger & Rockoff, 2010), or combinations of observation and student report (Brock et al., 2008; Kane et al., 2014; Raudenbush & Jean, 2014) the evidence indicates considerable consistency in identifying clusters of behaviors as reliably detectable and salient for student learning (Hamre et al., 2013). These common clusters include aspects of teachers' *social and emotional behaviors* toward students, their *practices related to classroom management*, and their *delivery of instruction* (Danielson, 2007; Marzano, 2014). Thus, although there is no standard lexicon for "effective teacher behaviors" – and the field lacks the precision and structure of a formal classification system – a scan of the evidence does converge on these common elements that serve as the conceptual foundation for the CLASS. The TTI Framework (Hamre & Pianta, 2007) draws heavily from earlier theoretical and empirical work in the educational and psychological literatures (e.g., Brophy, 1999; Eccles & Roeser, 2011) to describe an overarching theory of classroom practice, operationalized in the CLASS tool.

2.2 *The CLASS: Measuring and Describing the Quality of Teacher-Student Interactions*

As noted above and presented in Table 18.1, a key feature of the TTI framework is its *multi-level and nested* structure: teacher-student interaction is conceptualized and defined at multiple levels: domain, dimension, indicator, behavioral marker. At the most global, CLASS encodes teacher-student interaction within three broad domains—Emotional Support, Classroom Organization, and Instructional Support. At the next, more specific level, each domain is composed of a corresponding set of dimensions – teacher sensitivity, behavioral management, quality of feedback – which are the focus of the observation and for which the actual rating from low to high is obtained on a 1–7 scale. To inform those judgements and ratings, each dimension reflects a set of indicators that define the types of categories of behavior that correspond to that dimension. In this way, the CLASS and accompanying TTI Framework is like a classification system that defines the types of teacher behaviors that are salient for a broader feature of interactions. Finally, each indicator can be described in terms of its value or level of quality using specific behavioral markers that scale from low to high quality. The observer's job is to attend to and identify behavioral markers within the indicators for each dimension and make a judgement of the degree to which, as a collective pattern, these markers and indicators reflect a

Table 18.1 CLASS framework for early childhood and elementary classroom quality

Area	Dimension	Description
Emotional support	Positive climate	Reflects the overall emotional tone of the classroom and the connection between teachers and students.
	Negative climate	Reflects overall level of expressed negativity in the classroom between teachers and students (e.g., anger, aggression, irritability).
	Teacher sensitivity	Encompasses teachers' responsiveness to students' needs and awareness of students' level of academic and emotional functioning.
	Regard for student perspectives	The degree to which the teacher's interactions with students and classroom activities place an emphasis on students' interests, motivations, and points of view, rather than being very teacher-driven.
Classroom management	Behavior management	Encompasses teachers' ability to use effective methods to prevent and redirect misbehavior, by presenting clear behavioral expectations and minimizing time spent on behavioral issues.
	Productivity	Considers how well teachers manage instructional time and routines so that students have the maximum number of opportunity to learn.
	Instructional learning formats	The degree to which teachers maximize students' engagement and ability to learn by providing interesting activities, instruction, centers, and materials.
	Classroom chaos	The degree to which teachers ineffectively manage children in the classroom so that disruption and chaos predominate.
	Classroom management	The degree to which teachers provide clear instructions, rules, and routines that children clearly know and understand, as well as well-timed proactive behavioral strategies rather than control techniques.
	Child responsibility	The extent to which teachers provide children with the opportunity to take on roles and operate autonomously in the classroom.
Instructional support	Concept development	The degree to which instructional discussions and activities promote students' higher order thinking skills versus focus on rote and fact-based learning.
	Quality of feedback	Considers teachers' provision of feedback focused on expanding learning and understanding (formative evaluation), not correctness or the end product (summative evaluation).
	Language modeling	The quality and amount of teachers' use of language-stimulation and language-facilitation techniques during individual, small-group, and large-group interactions with children.
	Instructional conversation	Considers the extent to which teachers' verbal interactions with children are reciprocal and focus on the facilitation of reasoning, concept development, expression of ideas, and cognitive elaboration.
	Literacy instruction	The extent to which teachers reads to children, provides explicit phonics instruction, elaborates on books with comprehension and process questions, and exposes children to written language.
	Richness of instructional methods	The extent to which teacher use a variety of strategies to promote children's thinking and understanding of material at deeper and more complex level.

certain level of quality on that dimension. This multi-level framework is intentionally designed to yield scores that are more reflective of broad and organized patterns of teacher behavior while at the same time providing specific, concrete examples of use to observers and practitioners.

Research using the CLASS provides evidence confirming the three hypothesized common domains of teacher-child interactions in the TTI framework – Emotional Support, Classroom Organization, and Instructional Support – as a theoretically and empirically sound approach to describing teacher-student interactions in classrooms (Hamre et al., 2013). Results from a study of CLASS-derived observational data from over 4000 preschools to fifth grade U.S. classrooms (Hamre et al., 2013) supported the three-domain structure and analysis of CLASS-based observations in upper elementary and secondary grades from the Measures of Effective Teaching sample of more than 3000 classrooms (Kane & Staiger, 2012), also affirmed the importance of these three broad areas of practice. Thus, the evidence from large-scale use of CLASS observations in U.S. classrooms provides empirical support for the hypothesis of a common set of features on which teacher-student interactions can be described and distributed.

What do we know about the quality of interactions with teachers experienced by the typical American preschool or k-12 student? Many studies have found that quality of teacher-student interaction varies markedly across U.S. samples, ranging from sensitive and stimulating, to dismissive and harsh. In the National Center for Early Development and Learning's study of state prekindergarten programs, only 15 percent of classrooms demonstrated high-quality interactions in both emotional and instructional support, whereas 19 percent of classrooms scored well below the mean on almost all dimensions of emotional, organizational, and instructional supports (Pianta et al., 2005). Poor and African American children are more likely to experience less effective interactions in early childhood programs (Kuhfeld et al., 2019).

Evidence from national-level observations of American elementary school classrooms shows clearly that the nature and quality of the instructional and social supports offered to young students is generally low, and even lower for less advantaged students (NICHD ECCRN, 2005; Pianta et al., 2007; MET Project, 2010; Kane & Staiger, 2012). The Measures of Effective Teaching (MET) Study, funded by the Bill and Melinda Gates Foundation, reported on the nature of experiences across two consecutive years in more than 3000 4th-10th grade classrooms in 4 large school districts (Kane et al., 2014; Kane & Staiger, 2012). Using a suite of standardized observation protocols that scanned for general qualities of teachers' interactions toward students (including CLASS) and teaching practices relevant to specific content areas, the MET findings corroborate the impressions gleaned years earlier from the NICHD Study of Early Child Care and Youth Development observations – classroom learning experiences were largely rote in nature and rarely called for reasoning, problem solving, or analytic skills; instruction was delivered primarily in large groups; content was discrete and isolated rather than made relevant and connected to other knowledge; and students were engaged in very passive ways (Kane et al., 2014; Kane & Staiger, 2012).

2.3 *Teacher-Student Interactions and Student Outcomes*

In numerous studies, the three domains of teacher-student interactions described earlier (emotional, organization, instruction) have each been linked to students' social, emotional, regulatory, and cognitive development (see Downer et al., 2010 for a review). Effect sizes obtained between these ratings of the features of teachers' interactive behaviors and student outcomes such as achievement test scores are small (Brock et al., 2008; Burchinal et al., 2010; Mashburn et al., 2008; Pakarinen et al., 2011; Rimm-Kaufman et al., 2009), with larger correlations for students with higher risk profiles (Hamre & Pianta, 2005; McCartney et al., 2007), or for associations with students' motivation (Ferguson & Hirsch, 2014). In U.S. studies, children who come from low-income families, who are dual language learners, or who have problems with self-regulation appear to benefit even more from effective teacher-student interactions than do their more-resourced peers (e.g., Ansari et al., 2020; Desimone & Long, 2010; Hamre & Pianta, 2005). And children reap the most academic benefit from effective teacher-student interactions when they are exposed to such interactions for a number of years (Cash et al., 2018; Vernon-Feagans et al., 2019).

Although much of the research using classroom observation has been conducted in U.S. elementary classrooms, recent work in a variety of international settings—including Central and South America, Europe, and Asia—has also documented that teacher-child interactions support development and learning. For example, in a large-scale study of classroom quality and child outcomes in rural Ecuador that spanned the first two years of schooling (ages six and seven) in which children were assigned randomly to teachers, children's academic skills improved more when they were assigned to classrooms in which teachers demonstrated particularly high levels of instructional support (Campos et al., 2021). Other studies in Ecuador (Araujo et al., 2014), Chile (Yoshikawa et al., 2015), and Finland (Pakarinen et al., 2011), and from observations in secondary grades (Allen et al., 2011; Kane et al., 2014) have produced similar findings. Although the nature and magnitude of the associations between teacher-child interactions and student outcomes has varied across these studies, evidence is growing that elements of these interactions are important for children's learning across a wide spectrum of settings and cultures and perhaps a universal resource for children's development.

Most published studies have used statistical controls to reduce or adjust for *selection effects*—primarily, the concern that higher-achieving children may sort into classrooms in which teachers are more likely to display higher-quality interactions. However, evidence from recent intervention studies and random assignment studies demonstrates a more compelling causal link. For example, when teachers improve their practices after they receive training and coaching on teacher-student interactions, the children in their classrooms benefit academically, socially, and behaviorally (Pianta et al., 2021). Other evidence for a causal link between interactions and development comes from large-scale studies that randomly assigned children to classrooms to evaluate how classrooms affected achievement and

development. Two such studies have found significant associations between children's learning and their exposure to interactions (Campos et al., 2021; Yoshikawa et al., 2015). One of them, conducted in Ecuadorian first- and second-grade classrooms, estimated that teachers in the top 25 percent in terms of the quality of their interactions with students produced the equivalent of almost 9 months more of achievement growth among children than did teachers in the bottom 25 percent (Campos et al., 2021). Moreover, over the past 5–6 years several professional development interventions designed to improve teacher-student interaction – including a coaching model and a college course—provide additional empirical support for the unique value of teacher-student interactions by demonstrating positive impacts of targeted professional development on both teacher-student interaction and student outcomes, from preschool through high school (e.g., Allen et al., 2011; Boston Consulting Group, 2019; Pianta et al., 2020).

2.4 Summary of U.S. Findings

Across the available studies based on largely U.S. samples, we have presented a summary of findings concerning teacher-student interactions. By and large these findings suggest that features of teacher-student interactions are often described in terms of broad domains of emotional, organizational, and instructional behaviors, that can be measured reliably and at scale, using observational methods. The CLASS is one such example of an observational approach that has been used widely in the U.S. and studied in countries across the world. Numerous studies, mostly quasi-experimental in design but also including a small number of experiments (studies of students assigned randomly to teachers and teacher-focused intervention experiments), indicate that teacher-student interactions have a small and significant, and perhaps causal, impact on student outcomes. And finally, controlled evaluations demonstrate that teacher-student interactions are malleable and can be improved through focused feedback and improvements in teachers' knowledge and observational skills.

3 Method

3.1 Systematic Literature Search

To identify international research or education systems that used the CLASS, we completed a systematic search of published and unpublished literature, including several search engines (PsychInfo, ERIC, Google Scholar, Academic Search Complete, Education Research Complete, Education Full Text), databases for masters and dissertations (ProQuest and LIBRA Institutional Repository hosted out of

the University of Virginia), websites of documents from large scale studies. Citations were uploaded into Covidence software, where duplicates were removed, and the remaining entries were systematically screened. Journal articles, reports, briefs, or theses that include information about CLASS data from at least 20 lead or subject-specific teachers in preK-12 educational settings were retained. Thus, literature from toddlers or childcare settings, summer or after school programs, or that includes fewer than 20 teachers and/or does not include CLASS data in the document were excluded. Furthermore, in order to account for the quality of data collected, we excluded studies that did not include trained raters and that did not provide information about the reliability of CLASS observations. Finally, to ensure that our search was exhaustive, we emailed the first author from each document to request information about other published or unpublished documents that met our inclusion criteria and included any new documents in the database. The full database includes 365 documents from 133 studies, among which 52 published documents are from 19 studies that used the CLASS outside of the United States. The final international database includes 19 documents (all of which are peer reviewed journal articles from the 19 studies) that use the CLASS outside of the United States (see Fig. 18.1). All documents were coded for sample characteristics, CLASS data collection and analysis, CLASS data, and other study findings (see Table 18.2 for a selective overview).

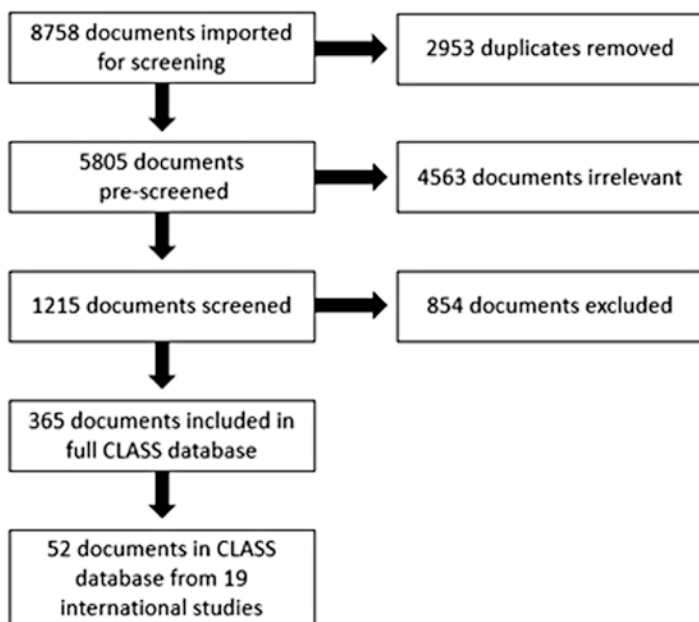


Fig. 18.1 PRISMA report of systematic search and screen of published and unpublished CLASS documents

Table 18.2 Studies of international classrooms that measure the quality of teacher-student interactions with the Classroom Assessment Scoring System

Description	Overall CLASS score						Emotional support mean			Instructional support			Classroom organization			
	# Teachers	# Students	Mean class size	Grade(s)	Mean	SD	ICC (Kappa)	Mean	S.D.	α	Mean	SD	α	Mean	S.D.	α
Niklas & Taylor, 2018	265 ^f	2123	9	Pre-K	4.04 ^a	0.93 ^a	0.80 ^a	5.14	0.91	0.87	2.38	0.96	0.85	4.60	0.92	0.89
Besnard & Letarte, 2017	53	180	3 ^a	Pre-K	4.22 ^a	0.69 ^a	0.80 ^a	5.26	0.72	.	2.59	0.63	.	4.82	0.73	.
Taut et al., 2019 ^e	51	1784	35	7
Yoshikawa et al., 2015	119	1876	21 ^a	Pre-K	3.55	.	(0.94)	4.64 ^a	.	.	1.72 ^a	.	.	4.30 ^a	.	.
Hu et al., 2016	180	5841 ^a	32	K	3.98 ^a	0.70 ^a	0.89 ^a	5.03	0.69	0.78	2.12	0.61	0.84	4.80	0.81	0.92
Slot et al., 2018	402	3132	21	Pre-K	4.66 ^a	0.48 ^a	0.90 ^a	5.85	0.42	0.73	2.45	0.55	0.64	5.69	0.47	0.83
Campos et al., 2021	2316	14,407	38	K-4	3.30	0.24	0.87	3.80	0.33	.	1.10	0.15	.	3.80	0.33	.
Pöysä et al., 2019	51	709	14 ^a	7	4.72 ^a	0.84 ^a	0.90 ^a	4.31 ^a	0.90 ^a	.	3.94 ^a	0.78 ^a	.	5.90 ^a	0.83 ^a	.
Pakarinen et al., 2010	49	679	14	K	4.81	0.79	0.85	5.13	0.80	0.93	3.97	0.92	0.88	5.34	0.66	0.90
Virtanen et al., 2018	46	949 ^a	21	6	4.73 ^a	0.65 ^a	0.55	5.04 ^a	0.66 ^a	0.83	3.01 ^a	0.70 ^a	0.82	4.00	0.57 ^a	0.90
Stuck et al., 2016	57 ^a	390	7 ^a	Pre-K	4.44 ^a	0.61 ^a	0.73	5.57 ^a	0.67 ^a	0.86	1.63 ^a	0.54 ^a	0.90	6.13 ^a	0.61 ^a	0.90
Von Suchodoletz et al., 2014	63	1323 ^a	21	Pre-K	4.21 ^a	0.85 ^a	0.80 ^b	5.33 ^a	0.75 ^a	0.89	2.47 ^a	0.78 ^a	0.81	4.82 ^a	1.02 ^a	0.85

The studies include data from 2186 prekindergarten and kindergarten classrooms, 2042 elementary school classrooms, and 177 secondary classrooms. For the CLASS observations, on average raters observed 3.3 cycles of classroom instruction over 1.6 days, about half of which were rated live (10/19), while the others rated video recordings of classroom interactions (9/19). Most of the studies describe their raters as being trained (18/19) and passing certification (15/19). The overall inter-rater reliability across studies (reported as intraclass correlations, percent agreement, or kappa scores) was reported as good to excellent, with the exception of two studies – one of Portuguese preschools (Cadima et al., 2014) and another of Finnish sixth grade classrooms (Virtanen et al., 2018), both of which had moderate inter-rater reliability (Ranganathan et al., 2017; Table 18.2).

4 Results

4.1 *Internal Consistency*

Reliability generalization reveals that the internal consistency of CLASS domains is sustained across the different cultural contexts. A reliability generalization is a meta-analytic technique that establishes 95% confidence intervals (Rodriguez & Maeda, 2006) for each of the three CLASS domains for the studies in which internal consistency coefficients were reported, which is mostly reported at the domain-level (see Table 18.2, Cohen's alpha, α). The Emotional Support domain had a reliability C.I. of 0.81 to 0.89, Instructional Support had a C.I. of 0.87 to 0.94, and Classroom Organization of 0.78 to 0.87. This indicates that the internal reliability for each domain was high across the international studies. This contributes important preliminary evidence that the TTI framework captures aspects of teacher-student interactions that are fundamental and appear in classrooms in very different cultural contexts.

4.2 *Factor Structure*

Several studies used the proposed 3 domain framework in which classroom quality consists of emotional support, instructional support, and classroom organization (Besnard & Letarte, 2017; Cadima et al., 2014; Castro et al., 2017; Gamlem & Munthe, 2014; Gasser et al., 2018; Niklas & Tayler, 2018; Pöysä et al., 2019; Sandstrom, 2012). Among the studies that evaluated the factor structure of the CLASS, support for 3-domain framework was found in early education classrooms across the globe, including prekindergarten samples in Chile (Yoshikawa et al., 2015 as cited in Leyva et al., 2015), Denmark (Slot et al., 2018), and Turkey (Ertürk Kara et al., 2017), and in kindergarten samples in Germany (Von Suchodoletz et al.,

2014), Vietnam (Hoang et al., 2018), and in China, where there was also support for a bi-factor model (Hu et al., 2016) (see papers for specific adjustments to factor analyses like correlating errors or residuals). One study of seventh graders in Chile (Taut et al., 2019) reported that they did not confirm the 3-factor structure and so instead chose to report the components of quality at the dimension-level (which we did not aggregate to the domain or overall levels of quality for meta-analysis or review).

In some cases, certain dimensions did not contribute to capturing classroom quality in a given cultural sample or setting. This is particularly the case with the Negative Climate dimension, which did not appear to be a significant component of the Emotional Support domain in several countries. In the first systematic examination of the CLASS in Europe, for example, Pakarinen et al. (2010) found that quality of the Finnish kindergarten teachers in their samples was best represented when the Negative Climate dimension was omitted.

Similarly, noting the poor discriminate validity of the Negative climate dimension in the previous study, Stuck et al. (2016) also omitted the dimension their study of 57 prekindergarten teachers in Germany. In another study of almost 180 prekindergarten teachers in Portugal, Cadima et al. (2018) found that when they omitted the Negative Climate dimension, the three-factor model provided the best relative fit to the data. It should be noted that contemporary guidance on the use of CLASS in research and in applied implementations suggests excluding Negative Climate from the domain-level computations.

Finally, in a study of sixth grade Finnish teachers, Virtanen et al. (2018) found support for a 3-factor model after excluding the Regard for Adolescent Perspectives and Instructional Learning Formats dimensions, each of which tended to cross-load with domains other than the hypothesized structure. These two dimensions have also been noted to cross-load in some U.S. studies (Hamre et al., 2013).

4.3 Levels of the Quality of Teacher-Student Interaction

Of considerable interest for this first multi-country view of teacher-student interaction was the pattern of levels of interaction quality seen across countries. Overall, the mean level of quality reported across the international studies reflects what we see in the American research: mostly mid (4) to middle-high scores (5) for the Emotional Support and the Classroom Organization domains, and mostly lower (2) to low-mid scores (3) for the Instructional Support domain (e.g., Harnes et al., 2014; La Paro et al., 2009). Internationally, the highest scores are reported in Classroom Organization, with multiple studies reporting a high score (mean level of almost or over 6), which is somewhat higher than in the U.S., in which the highest scores are typically associated with the Emotional Support domain, at least in younger-grade samples. Not dissimilar to results from the U.S., this multi-national analysis indicates the mean level of Instructional Support is 2.7 across the studies; several studies reported Instructional Support in the low range (1–2), with only a few reporting

mid-range scores (3–5). This pattern of low levels on the CLASS Instructional Support domain is consistent with U.S. findings and suggests that most of the instruction in classrooms has a focus on learning discrete facts and skills through instruction that has a rote focus.

To describe average quality across samples from each country, we generated means for the overall CLASS score that adjust for the reliability among raters in each study (Wiernik & Dahlke, 2020). Each overall CLASS mean reflects the average overall quality, within a range of error that in part relates to the level of alignment among raters. Adjusting for inter-reliability across samples provides a better sense of the range within which the true CLASS mean could reside. The corrected means account for inter-rater reliability by using the methods implemented in the psychmeta package (Dahlke & Wiernik, 2019; Wiernik & Dahlke, 2020). The two most common ways that reliability was reported in the selected studies were the intra-class coefficient (ICC) and percent agreement between raters. Overall, the quality of teacher-student interactions from these samples across the globe varies within the mid-range, with the overall mean adjusted for reliability at 3.69 (95% CI: 3.33, 4.06).

4.4 Teacher-Student Interaction and Student Outcomes

Due to variation in outcomes and outcome measures, it was not possible to use meta-analysis to assess how the quality of interactions measured with the CLASS relate to student outcomes. Instead, we review and synthesize the study findings in all documents across the studies.

Altogether, the international studies contribute to evidence that the quality of interactions with teachers shape children's developmental and academic success. In the first years of school, interaction quality promotes self-regulation among students in different cultural contexts. The overall quality of interactions is highly correlated with preschoolers' attention and impulse control in Turkey (Ertürk Kara et al., 2017), and cognitive self-regulation among socially disadvantaged preschoolers in Portugal (Cadima, Enrico, et al., 2016a). Furthermore, the Portuguese study suggests that teacher-student interactions can be a protective factor for young children at risk, such that interaction quality can be particularly effective in supporting students who are low in self-regulation skills (Cadima, Verschueren, et al., 2016b) and among children who are exposed to more family risk factors (Cadima, Enrico, et al., 2016a). Among kindergarten students in China, instructional support, in particular, is associated with growth in students' executive function skills (Hu et al., 2020). And in a large longitudinal experimental study of interaction quality in Ecuador, children in grades K-4 who were randomly assigned to teachers with higher quality interactions had higher executive function skills, particularly for working memory (Campos et al., 2021). Higher quality interactions also reduced the likelihood of behavioral problems in the same year (Campos et al., 2021).

Interactions that structure learning opportunities supports children's social development and adaptive classroom behavior in international settings. In a sample of Canadian preschoolers, Besnard and Letarte (2017) found that interactions that structure children's concept development and instructional learning support growth in social competence and overall adaptability, respectively. Similarly, among a sample of Finnish kindergarten students, the quality of instructional support was positively associated with empathy and negatively associated with disruptive behavior (Siekkinen et al., 2013) and less task avoidant behavior in class (Pakarinen et al., 2011). Furthermore, the quality of teachers' classroom organization predicted learning motivation among Finnish kindergartners (Pakarinen et al., 2010) and self-reports of behavioral and cognitive engagement among Finnish secondary students (Pöysä et al., 2019).

The international studies also verify that warm and supportive interactions with teachers are important to children throughout their education. Across various cultural settings, teachers' ability to identify and respond to the emotional needs of their students supported student engagement in learning. In Swedish preschools, emotional support predicted student engagement over time (Castro et al., 2017) and a combination of positive climate, instructional learning formats, and language modeling predicted children's engagement in literacy learning (Norling et al., 2015). In Finnish elementary classrooms, first graders who experienced low levels of emotional support were more likely to display passive avoidance when faced with academically challenging work in second grade (Pakarinen et al., 2014). Among Finnish adolescents, emotionally supportive interactions with teachers are associated with students' own report of their situational engagement (Pöysä et al., 2019). Emotional support also reflected and reinforced the quality of teachers' relationships with their students. In a sample of Swiss fifth graders, observer ratings of emotional support were related with students' perceptions of their teacher as caring and high level of emotional support protected students who were highly disengaged from academics from developing perceptions of their teacher as unjust (Gasser et al., 2018).

Each of the three domains of interaction quality are associated with direct assessment of academic skills across the various cultural contexts. Overall quality of interactions is associated with growth in both language and preliteracy skills among Danish preschoolers (Slot et al., 2018) and Ecuadorian K-fourth grade students, with the strongest effects in kindergarten and first grade (Campos et al., 2021). Researchers in the Ecuador study also found that the effects of experiencing high quality interactions with a kindergarten teacher are evident into sixth grade (Campos et al., 2021).

In early childhood education centers in Australia, the quality of teachers' instructional support predicted verbal abilities among children 4 years or older (Niklas & Tayler, 2018). In China, instructional support has been positively associated with reading, math, and science achievement among preschoolers (Hu et al., 2017) and emotional support has been linked with kindergartener's reading attitudes, and children with better reading attitudes benefited more from instructional support and exhibited greater gains in their vocabulary scores (Hu et al., 2018). Emotional

support in kindergarten was also positively associated with Finnish children's reading skills in first grade (Silinskas et al., 2017). In Portugal, the quality of teachers' classroom organization was positively associated with first grade students' vocabulary and print concepts, even after taking family risk and prior learning into account (Cadima et al., 2010).

There was also important evidence that interaction quality can address or exacerbate social disparities in education outcomes. In their study of Australian preschoolers, Niklas and Tayler (2018) found that, in classrooms with low quality interactions, the prestige of parents' occupations predicted children's verbal ability, whereas in high quality classrooms, there was no relationship between parent occupational prestige and verbal ability. Similarly, in classrooms with low quality organization, parent education predicted children's performance on mathematics assessments, whereas there was no relationship between parent education and mathematics achievement in classrooms rated high on classroom organization (Niklas & Tayler, 2018). Correspondingly, in a study of Portuguese students, Cadima et al. (2010) found that students with low math skills in preschool benefit more from high quality interactions with their first-grade teacher, which could contribute to narrowing math achievement gaps among students who start skills with disparate levels of math skills.

Together, research from international studies contributes additional empirical support for the teacher-student interactions as a developmentally salient feature of educational settings across the globe. In a combination of large-scale implementations, quasi-experimental, and experimental studies, the quality of teacher-student interactions predicts developmental and academic outcomes in very different cultural settings. The overall pattern of results suggests the value of teacher-student interactions for students' learning and development is significant and consistent across countries and cultures.

5 Conclusions and Discussion

In the educational context, teacher-student interactions play a fundamental role in determining the impact of teachers on student development and learning across wide-ranging countries and cultures. Describing, measuring, and improving teacher-student interactions are critical to large-scale efforts to build and improve public education systems.

The present study is an effort to draw upon theory and empirical research on teacher-student interaction conducted in the U.S. to examine the extent to which there is consistency in findings drawn from samples of teachers and students in non-U.S. countries across the globe.

By and large the results obtained from this multinational synthesis are notably consistent with those reported in U.S. samples. Across the 16 countries, 4400 teachers, and 42,000 students included in these analyses, empirical support was found for

the following conclusions: (1) teacher-student interactions can be describing using a common set of descriptors and reliably observed using those descriptors across countries that vary widely in cultural and educational circumstances; (2) teacher-student interactions appear to have a common underlying organization such that aspects of their emotional supports, instructional interactions, and classroom organization form a framework for description that can be used consistently across countries; (3) these three features of interaction have significant and beneficial impacts on students' learning and development.

Although not directly reported here and with many fewer exemplars internationally (e.g., Yoshikawa et al., 2015), it is clear from U.S. studies that these features of interaction can be improved through focused training and supports. Collectively, these are notable results with powerful implications for investments in workforce development systems that focus on teacher-student interaction as a means to improve the quality of educational opportunity and outcomes.

The conclusions above should be framed by certain caveats and limitations. The CLASS was used as a common classroom observation tool to capture general properties of classroom interactions, without modifications to reflect nuances unique to culture, ethnicity, race, or language. Moreover, the descriptive statistics reported (e.g., means, variance) are all drawn from convenience samples; none are representative of the countries' populations or school systems (this includes those from the U.S.). Therefore, cross-country comparisons in these indicators of effective teaching are not advised, nor is it appropriate to draw conclusions about the level of effective teaching in a given country. That said, the descriptive findings point to the potential use of observations, such as CLASS or other scalable measures, in samples more representative of countries or important political, geographic, or cultural groups, which might drive investments in education systems and teacher development.

With these general conclusions in mind, there are several implications for further research. Assuming the aim to use a common observational tool across countries, questions of interest might involve the extent to which characteristics of observers (e.g., prior knowledge, cultural background or differences, experience) are associated with differential levels of agreement. Additionally, questions related to training observers include whether observer reliability is related to the nature and amount of didactic training, practice in scoring video, and the types and ranges of video to be used in training. These questions essentially focus on the conditions that enable or limit the use of a common tool across wide-ranging cultures. Furthermore, even under circumstances in which a common tool might be applicable, research that informed refining both common and country/culture specific features of interaction that are important for students' learning and development, would inform observational systems that are best suited to a culture's uniqueness as well as capturing what common elements of effective teaching. Finally, research that helps to efficiently and cost-effectively scale measurement and improvement systems for teacher-student interaction will have considerable value for efforts to invest more systematically in improving public education systems across the globe.

References

- Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, *333*(6045), 1034–1037. <https://doi.org/10.1126/science.1207998>
- Ansari, A., Pianta, R. C., Whittaker, J. V., Vitiello, V. E., & Ruzek, E. A. (2020). Enrollment in public-prekindergarten and school readiness skills at kindergarten entry: Differential associations by home language, income, and program characteristics. *Early Childhood Research Quarterly*, *54*, 60–71. <https://doi.org/10.1016/j.ecresq.2020.07.011>
- Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2014). A helping hand? Teacher quality and learning outcomes in kindergarten. *Unpublished manuscript*.
- Besnard, T., & Letarte, M. J. (2017). Effect of male and female early childhood education teacher's educational practices on children's social adaptation. *Journal of Research in Childhood Education*, *31*(3), 453–464. <https://doi.org/10.1080/02568543.2017.1319445>
- Boston Consulting Group. (2019). *High quality classrooms lead to improved test scores*. Report to Dallas Independent School District.
- Brock, L. L., Nishida, T. K., Chiong, C., Grimm, K. J., & Rimm-Kaufman, S. E. (2008). Children's perceptions of the classroom environment and social and academic performance: A longitudinal analysis of the contribution of the responsive classroom approach. *Journal of School Psychology*, *46*(2), 129–149. <https://doi.org/10.1016/j.jsp.2007.02.004>
- Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of human development* (pp. 993–1028). John Wiley & Sons Inc.
- Brophy, J. (1999). Toward a model of the value aspects of motivation in education: Developing appreciation for. *Educational Psychologist*, *34*(2), 75–85. https://doi.org/10.1207/s15326985sep3402_1
- Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child-care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly*, *25*(2), 166–176. <https://doi.org/10.1016/j.ecresq.2009.10.004>
- Cadima, J., Aguiar, C., & Barata, M. C. (2018). Process quality in Portuguese preschool classrooms serving children at-risk of poverty and social exclusion and children with disabilities. *Early Childhood Research Quarterly*, *45*, 93–105. <https://doi.org/10.1016/j.ecresq.2018.06.007>
- Cadima, J., Enrico, M., Ferreira, T., Verschuere, K., Leal, T., & Matos, P. M. (2016a). Self-regulation in early childhood: The interplay between family risk, temperament and teacher-child interactions. *European Journal of Developmental Psychology*, *13*(3), 341–360. <https://doi.org/10.1080/17405629.2016.1161506>
- Cadima, J., Verschuere, K., Leal, T., & Guedes, C. (2016b). Classroom interactions, dyadic teacher-child relationships, and self-regulation in socially disadvantaged young children. *Journal of Abnormal Child Psychology*, *44*(1), 7–17. <https://doi.org/10.1007/s10802-015-0060-5>
- Cadima, J., Peixoto, C., & Leal, T. (2014). Observed classroom quality in first grade: Associations with teacher, classroom, and school characteristics. *European Journal of Psychology of Education*, *29*(1), 139–158. <https://doi.org/10.1007/s10212-013-0191-4>
- Cadima, J., Leal, T., & Burchinal, M. (2010). The quality of teacher-student interactions: Associations with first graders' academic and behavioral outcomes. *Journal of School Psychology*, *48*(6), 457–482. <https://doi.org/10.1016/j.jsp.2010.09.001>
- Campos, A., Carneiro, P., Cruz-Aguayo, Y., & Norbert, S. (2021). *Interactions: Do teacher behaviors predict achievement, executive function, and non-cognitive outcomes in elementary school?*
- Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2019). *Experimental estimates of education production functions: Sensitive periods and dynamic complementarity*. Institute for Fiscal Studies, University College London.

- Cash, A. H., Ansari, A., Grimm, K. J., & Pianta, R. C. (2018). Power of two: The impact of 2 years of high-quality teacher child interactions. *Early Education and Development, 30*(1), 60–81. <https://doi.org/10.1080/10409289.2018.1535153>
- Castro, S., Granlund, M., & Almqvist, L. (2017). The relationship between classroom quality-related variables and engagement levels in Swedish preschool classrooms: A longitudinal study. *European Early Childhood Education Research Journal, 25*(1), 122–135. <https://doi.org/10.1080/1350293X.2015.1102413>
- Dahlke, J. A., & Wiernik, B. M. (2019). Psychmeta: An R package for psychometric meta-analysis. *Applied Psychological Measurement, 43*(5), 415–416. <https://doi.org/10.1177/0146621618795933>
- Danielson, C. (2007). *Enhancing professional practice: A framework for teaching*. ASCD.
- Desimone, L. M., & Long, D. (2010). Teacher effects and the achievement gap: Do teacher and teaching quality influence the achievement gap between black and white and high- and low-SES students in the early grades? *Teachers College Record, 112*(12), 3024–3073.
- Downer, J., Sabol, T. J., & Hamre, B. (2010). Teacher–child interactions in the classroom: Toward a theory of within-and cross-domain links to children’s developmental outcomes. *Early Education and Development, 21*(5), 699–723. <https://doi.org/10.1080/10409289.2010.497453>
- Eccles, J. S., & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence, 21*(1), 225–241. <https://doi.org/10.1111/j.1532-7795.2010.00725.x>
- Ertürk Kara, H. G., Gönen, M. S., & Pianta, R. (2017). The examination of the relationship between the quality of teacher-child interaction and children’s self-regulation skills. *H.U. Journal of Education, 32*(4), 880–895.
- Ferguson, R. F., with Hirsch, E. (2014). How working conditions predict teaching quality and student outcomes. In T. J. Kane, K. A. Kerr, & R. C. Pianta (Eds.), *Designing teaching evaluation systems* (pp. 332–380). Wiley.
- Gamlem, S. M., & Munthe, E. (2014). Mapping the quality of feedback to support students’ learning in lower secondary classrooms. *Cambridge Journal of Education, 44*(1), 75–92. <https://doi.org/10.1080/0305764X.2013.855171>
- Gasser, L., Grütter, J., Buholzer, A., & Wettstein, A. (2018). Emotionally supportive classroom interactions and students’ perceptions of their teachers as caring and just. *Learning and Instruction, 54*, 82–92. <https://doi.org/10.1016/j.learninstruc.2017.08.003>
- Goe, L. (2007). The link between teacher quality and student outcomes: A research synthesis. *National comprehensive center for teacher quality*.
- Grossman, P., Cohen, J., Ronfeldt, M., & Brown, L. (2014). The test matters: The relationship between classroom observation scores and teacher value added on multiple types of assessment. *Educational Researcher, 43*(6), 293–303. <https://doi.org/10.3102/0013189X14544542>
- Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S. M., et al. (2013). Teaching through interactions: Testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *The Elementary School Journal, 113*(4), 461–487. <https://doi.org/10.1086/669616>
- Hamre, B. K., & Pianta, R. C. (2007). Learning opportunities in preschool and early elementary classrooms. In R. C. Pianta, M. J. Cox, & K. L. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 49–83). Paul H Brookes Publishing.
- Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development, 76*(5), 949–967. <https://doi.org/10.1111/j.1467-8624.2005.00889.x>
- Harnes, B., Hatfield, B., Pianta, R., & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher-child interactions: Associations with preschool children’s development. *Child Development, 85*(3), 1257–1274. <https://doi.org/10.1111/cdev.12184>
- Heckman, J. J. (2000). Policies to foster human capital. *Research in Economics, 54*, 3–56. <https://doi.org/10.1006/reec.1999.0225>

- Hill, S. (2010). The millennium generation: Teacher-researchers exploring new forms of literacy. *Journal of Early Childhood Literacy*, 10(3), 314–340. <https://doi.org/10.1177/1468798410372820>
- Hoang, N., Holopainen, L., & Siekkinen, M. (2018). Quality of teacher–child interactions and its relations to children’s classroom engagement and disaffection in Vietnamese kindergartens. *International Journal of Early Years Education*, 26(4), 387–402. <https://doi.org/10.1080/09669760.2018.1478281>
- Hu, B. Y., Fan, X., Gu, C., & Yang, N. (2016). Applicability of the classroom assessment scoring system in Chinese preschools based on psychometric evidence. *Early Education and Development*, 27(5), 714–734. <https://doi.org/10.1080/10409289.2016.1113069>
- Hu, B. Y., Zhou, Y., Chen, L., Fan, X., & Winsler, A. (2017). Preschool expenditures and Chinese children’s academic performance: The mediating effect of teacher-child interaction quality. *Early Childhood Research Quarterly*, 41, 37–49. <https://doi.org/10.1016/j.ecresq.2017.05.002>
- Hu, B. Y., Wu, H., Curby, T. W., Wu, Z., & Zhang, X. (2018). Teacher–child interaction quality, attitudes toward reading, and literacy achievement of Chinese preschool children: Mediation and moderation analysis. *Learning and Individual Differences*, 68, 1–11. <https://doi.org/10.1016/j.lindif.2018.09.004>
- Hu, B. Y., Fan, X., Wu, Y., LoCasale-Crouch, J., & Song, Z. (2020). Teacher–child interaction quality and Chinese children’s academic and cognitive development: New perspectives from piecewise growth modeling. *Early Childhood Research Quarterly*, 51, 242–255. <https://doi.org/10.1016/j.ecresq.2019.10.003>
- Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Research Paper. MET Project. Bill & Melinda Gates Foundation.
- Kane, T., Kerr, K., & Pianta, R. (2014). *Designing teacher evaluation systems: New guidance from the measures of effective teaching project*. John Wiley & Sons.
- Kuhfeld, M., Condron, D. J., & Downey, D. B. (2019). When does inequality grow? A seasonal analysis of racial/ethnic disparities in learning from kindergarten through eighth grade. *Educational Researcher*, 50(4), 225–238. <https://doi.org/10.3102/0013189X20977854>
- La Paro, K. M., Siepak, K., & Scott-Little, C. (2009). Assessing beliefs of preservice early childhood education teachers using Q-sort methodology. *Journal of Early Childhood Teacher Education*, 30(1), 22–36. <https://doi.org/10.1080/10901020802667805>
- Lemov, D. (2010). *Teach like a champion: 49 techniques that put students on the path to college (K-12)*. John Wiley & Sons.
- Leyva, D., Weiland, C., Barata, M., Yoshikawa, H., Snow, C., Treviño, E., & Rolla, A. (2015). Teacher–child interactions in Chile and their associations with prekindergarten outcomes. *Child Development*, 86(3), 781–799. <https://doi.org/10.1111/cdev.12342>
- Loeb, S., Kalogrides, D., & Bételle, T. (2012). Effective schools: Teacher hiring, assignment, development, and retention. *Education Finance and Policy*, 7(3), 269–304. https://doi.org/10.1162/EDFP_a_00068
- Marzano, R. J. (2014). Proximal versus distal validity coefficients for teacher observational instruments. *The Teacher Educator*, 49(2), 89–96. <https://doi.org/10.1080/08878730.2014.885783>
- Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., et al. (2008). Measures of classroom quality in prekindergarten and children’s development of academic, language, and social skills. *Child Development*, 79(3), 732–749. <https://doi.org/10.1111/j.1467-8624.2008.01154.x>
- Maulana, R., André, S., Helms-Lorenz, M., Ko, J., Chun, S., Shahzad, A., et al. (2021). Observed teaching behaviour in secondary education across six countries: Measurement invariance and indication of cross-national variations. *School Effectiveness and School Improvement*, 32(1), 64–95. <https://doi.org/10.1080/09243453.2020.1777170>
- McCartney, K., Dearing, E., Taylor, B. A., & Bub, K. L. (2007). Quality childcare supports the achievement of low-income children: Direct and indirect pathways through caregiving and the home environment. *Journal of Applied Developmental Psychology*, 28(5–6), 411–426. <https://doi.org/10.1016/j.appdev.2007.06.010>

- Morrison, F. J., & Connor, C. M. (2002). Understanding schooling effects on early literacy: A working research strategy. *Journal of School Psychology, 40*(6), 493–500.
- National Institute of Child Health and Human Development Early Child Care Research Network (2005). Early childcare and children's development in the primary grades. *American Educational Research Journal, 42*, 537–570. <https://doi.org/10.3102/00028312042003537>
- Niklas, F., & Tayler, C. (2018). Room quality and composition matters: Children's verbal and numeracy abilities in Australian early childhood settings. *Learning and Instruction, 54*, 114–124. <https://doi.org/10.1016/j.learninstruc.2017.08.006>
- Norling, M., Sandberg, A., & Almqvist, L. (2015). Engagement and emergent literacy practices in Swedish preschools. *European Early Childhood Education Research Journal, 23*(5), 619–634. <https://doi.org/10.1080/1350293X.2014.996423>
- Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*(3), 237–257. <https://doi.org/10.3102/01623737026003237>
- Pakarinen, E., Aunola, K., Kiuru, N., Lerkkanen, M. K., Poikkeus, A. M., Siekkinen, M., & Nurmi, J. E. (2014). The cross-lagged associations between classroom interactions and children's achievement behaviors. *Contemporary Educational Psychology, 39*(3), 248–261. <https://doi.org/10.1016/j.cedpsych.2014.06.001>
- Pakarinen, E., Kiuru, N., Lerkkanen, M. K., Poikkeus, A. M., Ahonen, T., & Nurmi, J. E. (2011). Instructional support predicts children's task avoidance in kindergarten. *Early Childhood Research Quarterly, 26*(3), 376–386. <https://doi.org/10.1016/j.ecresq.2010.11.003>
- Pakarinen, E., Lerkkanen, M. K., Poikkeus, A. M., Kiuru, N., Siekkinen, M., Rasku-Puttonen, H., & Nurmi, J. E. (2010). A validation of the classroom assessment scoring system in Finnish kindergartens. *Early Education and Development, 21*(1), 95–124. <https://doi.org/10.1080/10409280902858764>
- Patrick, H., Anderman, L. H., Ryan, A. M., Edelin, K. C., & Midgley, C. (2001). Teachers' communication of goal orientations in four fifth-grade classrooms. *The Elementary School Journal, 102*(1), 35–58. <https://doi.org/10.1086/499692>
- Pianta, R. C. (1999). Enhancing relationships between children and teachers. *American Psychological Association*. <https://doi.org/10.1037/10314-000>
- Pianta, R. C. (2016). Teacher–Student interactions: Measurement, impacts, improvement, and policy. *Policy Insights from the Behavioral and Brain Sciences, 3*(1), 98–105. <https://doi.org/10.1177/2372732215622457>
- Pianta, R. C., Hamre, B. K., & Nguyen, T. (2020). Measuring and improving quality in early care and education. *Early Childhood Research Quarterly, 51*, 285–287. <https://doi.org/10.1016/j.ecresq.2019.10.013>
- Pianta, R. C., & Allen, J. P. (2008). Building capacity for positive youth development in secondary school classrooms: Changing teachers' interactions with students. In M. B. Shinn & H. Yoshikawa (Eds.), *Toward positive youth development: Transforming schools and community programs* (pp. 21–40). Oxford University Press.
- Pianta, R. C., Lipscomb, D., & Ruzek, E. (2021). *Coaching of teachers to improve students' school readiness skills: Indirect effects of teacher-student interaction* (Vol. 92, pp. 2509–2528). In press.
- Pianta, R. C., Mashburn, A., Downer, J., Hamre, B., & Justice, L. (2008). Effects of web-mediated professional development resources on teacher-child interactions in pre-kindergarten classrooms. *Early Childhood Research Quarterly, 23*(4), 431–451. <https://doi.org/10.1016/j.ecresq.2008.02.001>
- Pianta, R. C., Belsky, J., Houts, R., & Morrison, F. (2007). Opportunities to learn in America's elementary classrooms. *Science, 315*(5820), 1795–1796. <https://doi.org/10.1126/science.1139719>
- Pianta, R. C., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science, 9*(3), 144–159. https://doi.org/10.1207/s1532480xads0903_2

- Pöysä, S., Vasalampi, K., Muotka, J., Lerkkanen, M. K., Poikkeus, A. M., & Nurmi, J. E. (2019). Teacher–student interaction and lower secondary school students’ situational engagement. *British Journal of Educational Psychology*, *89*(2), 374–392. <https://doi.org/10.1111/bjep.12244>
- Ranganathan, P., Pramesh, C. S., & Aggarwal, R. (2017). Common pitfalls in statistical analysis: Measures of agreement. *Perspectives in Clinical Research*, *8*(4), 187–191.
- Raudenbush, S. W., & Jean, M. (2014). To what extent do student perceptions of classroom quality predict teacher value added? In T. J. Kane, K. A. Kerr, & R. C. Pianta (Eds.), *Designing teacher evaluation systems* (pp. 170–201). Jossey Bass.
- Reardon, S. F., Valentino, R. A., Kalogrides, D., Shores, K. A., & Greenberg, E. H. (2013). *Patterns and trends in racial academic achievement gaps among states, 1999–2011*. Center for Education Policy Analysis, Stanford University. Retrieved from <https://cepa.stanford.edu/content/patterns-and-trends-racial-academic-achievement-gaps-among-states-1999-2011>.
- Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., Nathanson, L., & Brock, L. L. (2009). The contribution of children’s self-regulation and classroom quality to children’s adaptive behaviors in the kindergarten classroom. *Developmental Psychology*, *45*(4), 958–972. <https://doi.org/10.1037/a0015861>
- Rodriguez, M. C., & Maeda, Y. (2006). Meta-analysis of coefficient alpha. *Psychological Methods*, *11*(3), 306–322.
- Ruzek, E. A., Hafen, C. A., Allen, J. P., Gregory, A., Mikami, A. Y., & Pianta, R. C. (2016). How teacher emotional support motivates students: The mediating roles of perceived peer relatedness, autonomy support, and competence. *Learning & Instruction*, *42*, 95–103. <https://doi.org/10.1016/j.learninstruc.2016.01.004>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78.
- Sabol, T. J., Hong, S. S., Pianta, R. C., & Burchinal, M. R. (2013). Can rating pre-k programs predict children’s learning? *Science*, *341*(6148), 845–846. <https://doi.org/10.1126/science.1233517>
- Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement [Research report]*. University of Tennessee Value-Added Research and Assessment Center.
- Sandstrom, H. (2012). The characteristics and quality of pre-school education in Spain. *International Journal of Early Years Education*, *20*(2), 130–158.
- Siekkinen, M., Pakarinen, E., Lerkkanen, M. K., Poikkeus, A. M., Salminen, J., Poskiparta, E., & Nurmi, J. E. (2013). Social competence among 6-year-old children and classroom instructional support and teacher stress. *Early Education & Development*, *24*(6), 877–897. <https://doi.org/10.1080/10409289.2013.745183>
- Silinskas, G., Pakarinen, E., Lerkkanen, M. K., Poikkeus, A. M., & Nurmi, J. E. (2017). Classroom interaction and literacy activities in kindergarten: Longitudinal links to grade 1 readers at risk and not at risk of reading difficulties. *Contemporary Educational Psychology*, *51*, 321–335. <https://doi.org/10.1016/j.cedpsych.2017.09.002>
- Slot, P. L., Bleses, D., Justice, L. M., Markussen-Brown, J., & Højen, A. (2018). Structural and process quality of Danish preschools: Direct and indirect associations with children’s growth in language and preliteracy skills. *Early Education and Development*, *29*(4), 581–602. <https://doi.org/10.1080/10409289.2018.1452494>
- Staiger, D. O., & Rockoff, J. E. (2010). Searching for effective teachers with imperfect information. *Journal of Economic Perspectives*, *24*(3), 97–118. <https://doi.org/10.1257/jep.24.3.97>
- Stuck, A., Kammermeyer, G., & Roux, S. (2016). The reliability and structure of the classroom assessment scoring system in German pre-schools. *European Early Childhood Education Research Journal*, *24*(6), 873–894. <https://doi.org/10.1080/1350293X.2016.1239324>
- Taut, S., Jiménez, D., Puente-Duran, S., Palacios, D., Godoy, M. I., & Manzi, J. (2019). Evaluating the quality of teaching: Can there be valid differentiation in the middle of the performance distribution? *School Effectiveness and School Improvement*, *30*(3), 328–348. <https://doi.org/10.1080/09243453.2018.1510842>

- Van Hover, S., Hicks, D., & Cotton, S. (2012). "Can you Make 'Historiography' sound more friendly?": Towards the construction of a reliable and validated history teaching observation instrument. *The History Teacher*, 45(4), 603–612.
- Vernon-Feagans, L., Mokrova, I. L., Carr, R. C., Garrett-Peters, P. T., Burchinal, M., & Family Life Project Key Investigators. (2019). Cumulative years of classroom quality from kindergarten to third grade: Prediction to children's third grade literacy skills. *Early Childhood Research Quarterly*, 47, 531–540. <https://doi.org/10.1016/j.ecresq.2018.06.005>
- Virtanen, T. E., Pakarinen, E., Lerkkanen, M. K., Poikkeus, A. M., Siekkinen, M., & Nurmi, J. E. (2018). A validation study of classroom assessment scoring system—Secondary in the Finnish school context. *The Journal of Early Adolescence*, 38(6), 849–880. <https://doi.org/10.1177/0272431617699944>
- Von Suchodoletz, A., Fäsche, A., Gunzenhauser, C., & Hamre, B. K. (2014). A typical morning in preschool: Observations of teacher–child interactions in German preschools. *Early Childhood Research Quarterly*, 29(4), 509–519. <https://doi.org/10.1016/j.ecresq.2014.05.010>
- Wang, M. T., Hofkens, T. L., & Ye, F. (2020). Classroom quality and adolescent learning in mathematics: A multi-informant perspective. *Journal of Youth and Adolescence*, 49, 1987, 2002. Advance online publication. doi:<https://doi.org/10.1007/210964-020-01195-0>.
- Wang, M. T., & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement, and academic achievement in middle school. *American Educational Research Journal*, 47(3), 633–662. <https://doi.org/10.3102/0002831209361209>
- Wiernik, B. M., & Dahlke, J. A. (2020). Obtaining unbiased results in meta-analysis: The importance of correcting for statistical artifacts. *Advances in Methods and Practices in Psychological Science*, 3(1), 94–123. <https://doi.org/10.1177/2515245919885611>
- Yoshikawa, H., Leyva, D., Snow, C. E., Treviño, E., Barata, M., Weiland, C., et al. (2015). Experimental impacts of a teacher professional development program in Chile on preschool classroom quality and child outcomes. *Developmental Psychology*, 51(3), 309–322. <https://doi.org/10.1037/a0038785>

Tara Hofkens, PhD, is a Research Assistant Professor at the University of Virginia School of Education and Human Development.

Robert C. Pianta, PhD, is Dean of the UVA School of Education and Human Development, Novartis US Foundation Professor of Education, Professor of Psychology, and founding director of the Center for Advanced Study of Teaching and Learning at the University of Virginia. Dr. Pianta's research and policy interests focus on the intersection of education and human development. In particular his work has advanced conceptualization and measurement of teacher-student relationships and documented their contributions to students' learning and development. Dr. Pianta has led research and development on measurement tool and interventions that help teachers interact with students more effectively and that are used widely in the United States and around the world. Dr. Pianta received a BS and an MA in Special Education from the University of Connecticut and a PhD in Psychology from the University of Minnesota.

Bridget Hamre, PhD, is a Research Associate Professor at the University of Virginia School of Education and Human Development. Dr. Hamre's areas of expertise include student-teacher relationships and classroom processes that promote positive academic and social development for young children. She is deeply committed to working with education leaders to help bridge the "research to practice" divide. She is also currently serving as CEO at Teachstone, an organization founded to support partners in use of CLASS at federal, state, and local levels in the U.S. and internationally.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 19

Affective Student–Teacher Relationships and Students’ Engagement: A Cross–Cultural Comparison of China and The Netherlands



Debora Roorda, Mengdi Chen, and Marjolein Zee

Abstract Ample evidence has been found for the association between affective, dyadic student–teacher relationships and students’ engagement with schoolwork in Western, individualistic countries. There are far fewer studies, however, examining this association in Eastern, collectivistic countries. As maintaining harmony in interpersonal relationships plays a crucial role in collectivistic countries, student–teacher relationships may even be more important in collectivistic countries than in individualistic countries. In the present study, we therefore investigated cross–cultural differences in the strength of associations between student–teacher relationship quality and students’ engagement based on data from the Netherlands (a Western country) and China (an Eastern country). The Dutch sample included 789 students (51.1% girls) and the Chinese sample included 588 students (52.9% girls) from grades 3 to 6 of elementary school. Students reported about the quality of their relationship with their teacher (closeness, conflict) and their behavioral and emotional engagement with schoolwork. Hierarchical linear modeling showed that the positive association between closeness and both behavioral and emotional engagement was stronger for the Chinese sample than for the Dutch sample. In contrast, the negative association between conflict and both behavioral and emotional engagement did not differ across countries. To conclude, closeness may be more relevant for Chinese students’ engagement than would be expected based on Western studies,

D. Roorda (✉)

Research Institute of Child Development and Education, University of Amsterdam,
Amsterdam, Netherlands
e-mail: D.L.Roorda@uva.nl

M. Chen

Faculty of Education, University of Macau, Macau, China

M. Zee

Department of Psychology, Education and Child Studies, Erasmus University Rotterdam,
Rotterdam, The Netherlands

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_19

423

whereas conflict seems to be equally harmful in both cultures. Therefore, developing relationship-focused interventions for Chinese teachers and students seems important, either by adapting Western programs or by developing new programs especially designed for Chinese schools.

Keywords Affective teacher–student relationships · Behavioral engagement · Emotional engagement · Cross–cultural comparison · Upper elementary students

1 Affective Student–Teacher Relationships and Students’ Engagement: Differences Between China and the Netherlands

Previous research has generated convincing evidence that the emotional bond between teachers and individual students (i.e., affective quality of dyadic student–teacher relationships) affects elementary students’ school adjustment, such as their engagement with schoolwork (e.g., Archambault et al., 2013; Hamre & Pianta, 2001; Hughes, 2011). Most of these studies, however, were conducted in Western, individualistic countries, whereas this topic remains relatively understudied in Eastern, collectivistic countries. Some evidence has been found that observed teacher–student interactions are associated with students’ school adjustment in Eastern, collectivistic countries as well (e.g., Hu et al., 2017, 2021; Hoang et al., 2018). However, these studies focused on interactions between teachers and groups of students (i.e., teacher style or classroom climate) and not on dyadic relationships, which are the focus of the present study.

As maintaining harmonious relationships with significant others plays a central role in collectivistic cultures (Triandis, 2018), the impact of student–teacher relationships on students’ engagement with schoolwork may even be larger in collectivistic cultures than in individualistic cultures. Still, there is a lack of studies comparing the strength of associations between dyadic student–teacher relationships and students’ engagement with schoolwork across different countries. The present study therefore used data from both the Netherlands (a Western, individualistic country) and China (an Eastern, collectivistic country) to examine the existence of potential cross–cultural differences in the strength of associations between student–teacher relationships and engagement.

2 Student–Teacher Relationships and Students’ Engagement with Schoolwork

Research focusing on the affective quality of dyadic student–teacher relationships is often based on attachment theory (Pianta, 1999; Verschueren & Koomen, 2012). According to this theory, student–teacher relationships high in closeness (i.e., the

degree of warmth, open communication, and trust in the relationship) help students feel emotionally secure. Emotional security, in turn, is considered a necessary precondition for students’ optimal exploration of the classroom environment and for being engaged with schoolwork. In contrast, student–teacher relationships characterized by high levels of conflict (i.e., the level of negativity, tension, and hostility in the relationship) will hamper students’ emotional security and, hence, limit their engagement with schoolwork (Verschuere & Koomen, 2012). Engagement refers to students’ participation in schoolwork (i.e., behavioral engagement, such as effort, persistence, and concentration) as well as their feelings and emotions toward schoolwork (i.e., emotional engagement, such as enjoyment, satisfaction, and boredom; Skinner et al., 2009).

Studies conducted in Western countries (i.e., countries in North America, Northwestern Europe, and Australia) found ample evidence for the hypothesized association between affective student–teacher relationships and students’ engagement with schoolwork. For example, Zee and Koomen (2019) showed that student–teacher closeness was associated with more behavioral and emotional engagement in upper elementary students over time. A meta-analytic study based on 189 studies also revealed that positive student–teacher relationships (e.g., closeness) were associated with higher engagement with schoolwork (including both behavioral and emotional aspects). In contrast, negative relationships (e.g., conflict) were associated with less engagement (Roorda et al., 2017). Moreover, the same associations were found in a subsample including longitudinal studies only, indicating that associations between student–teacher relationship quality and engagement hold over time (Roorda et al., 2017). However, most of these studies were conducted in the United States of America (USA; $k = 111$) or other Western countries ($k = 50$), such as Belgium, Germany, the Netherlands, Norway, Canada, and Australia, and cultural differences in the strength of associations were not investigated.

3 Cultural Differences in Associations Between Student–Teacher Relationships and Engagement

According to the developmental systems model (Pianta et al., 2003), cultural values play an important role in the development of student–teacher relationships and their impact on students’ school adjustment. With regard to cultural values, a distinction is often made between individualistic cultures and collectivistic cultures (Hofstede et al., 2010; Triandis, 2001, 2018; Triandis et al., 1988). In individualistic cultures, ties between individuals tend to be loose and people are usually relatively independent from their in-groups (e.g., family, tribe, nation). In such cultures, personal autonomy is especially valued and it can be considered shameful to depend too much on others. People are expected to fulfill their own needs and usually base their behaviors and decisions on their own goals and values. In contrast, in collectivistic cultures, interpersonal interdependence is high, with ties between individuals being

strong and people being inclined to depend much on their in-groups. In such cultures, group loyalty is highly valued and working as a group and supporting others is essential. Common goals are considered more important than desires of individuals and people tend to base their decisions and behaviors on norms and values of significant others (Hofstede et al., 2010; Triandis, 2001, 2018; Triandis et al., 1988). Furthermore, values as respect and obedience to authority figures (e.g., teachers) are important in collectivistic cultures and students are also inclined to admire their teachers more than in individualistic cultures (Li, 2010; Triandis, 2018). Due to the higher degree of interpersonal interdependency and the importance of harmonious relationships in collectivistic cultures, relationships with teachers may have a larger impact on students' engaged behaviors and emotions in Eastern, collectivistic countries than in Western, individualistic countries.

In line with this idea, Zhou et al. (2012) found that relatedness with the teacher was positively associated with students' behavioral engagement in China but not in the USA. Likewise, a meta-analysis based on 65 studies (including 12 Asian studies) revealed that the association between teacher support and students' negative academic emotions (i.e., indicator of emotional disengagement) was stronger for East-Asian students than for Western-European and American students (Lei et al., 2018). In contrast, the association between teacher support and positive academic emotions appeared to be stronger in Western-European and American samples than in East-Asian samples (Lei et al., 2018).

To solve this inconsistency in findings, more research on cross-cultural differences in associations between dyadic student-teacher relationships and students' engagement seems to be needed. Furthermore, Lei et al. (2018) and Zhou et al. (2012) did not examine the impact of negative relationship dimensions (e.g., conflict), whereas previous research suggests that negative student-teacher relationships are more influential for elementary students' engagement with schoolwork than positive relationship dimensions (see Roorda et al., 2011, for a meta-analysis).

From a cross-cultural perspective, negative relationship dimensions are also interesting to study, as there tends to be a larger power distance and more respect for authority in schools in collectivistic countries than in individualistic countries (Hofstede et al., 2010; Li, 2010). In schools with a large power distance, students usually treat teachers with respect and deference and it is not appreciated if students publicly contradict or criticize their teachers. In schools in individualistic countries, however, teachers usually treat their students more as equals and arguing and disagreeing with teachers is more commonly accepted (Hofstede et al., 2010). Due to the larger power distance in collectivistic cultures, students may be more sensitive to and more frightened by conflictual relationships with teachers. As such, high levels of student-teacher conflict may even be more harmful for students' engagement in Eastern, collectivistic countries than in Western, individualistic countries. Therefore, the present cross-cultural comparison not only included closeness as relationship dimension but also focused on student-teacher conflict.

4 The Present Study

In the present study, we investigated the extent to which there are cultural differences in the strength of associations between student–teacher closeness and conflict and students’ behavioral and emotional engagement with schoolwork. In doing so, we focused on a sample of third to sixth graders from China (an Eastern, collectivistic country) and the Netherlands (a Western, individualistic country). Apart from logistical reasons, China and the Netherlands are interesting to compare, because of their distinct differences on individualism (i.e., the extent of interdependence amongst members of a society) and power distance (i.e., the degree to which a society believes that inequalities amongst people are acceptable; Hofstede et al., 2010). More specifically, in the Netherlands, independence of individuals is highly valued (score of 80 on individualism on a scale from 1 to 120; Hofstede Insights, n.d.), whereas large power differences among people are less accepted (score of 38 on power distance). In contrast, the Chinese society generally values interdependence among people (score of 20 on individualism) and generally accepts power differences between people (score of 80 on power distance; Hofstede Insights, n.d.). These societal values are considered to influence daily interactions and relationships between teachers and students and their impact on students’ school adjustment (Chen et al., 2019; Hofstede et al., 2010; Pianta et al., 2003).

We hypothesized that closeness would be positively associated with students’ behavioral and emotional engagement, whereas conflict would be negatively associated with behavioral and emotional engagement (Roorda et al., 2017; Zee & Koomen, 2019). Based on the higher interpersonal interdependence, the larger power distance, and the larger respect for authority in collectivistic countries (Hofstede et al., 2010; Li, 2010; Triandis, 2001, 2018), we expected that these associations would be stronger in the Chinese sample than in the Dutch sample.

5 Methods

5.1 Participants

The Dutch sample consisted of 789 students (51.1% girls) from 35 classrooms from eight regular elementary schools. The Chinese sample included 588 students (52.9% girls) from 14 classrooms from three regular elementary schools. In both samples, students were in third to sixth grade. However, as formal education starts 1 year later in China than in most Western countries, students in the Chinese sample ($M_{\text{age}} = 11.49$ years, $SD = 1.29$; range = 9–14 years) were somewhat older than in the Dutch sample ($M_{\text{age}} = 9.99$ years, $SD = 1.24$; range = 7–13 years; $t(1192.48) = -21.50, p < .001$). Furthermore, the number of students per classrooms

was higher in China ($M_{\text{classroom size}} = 43$ students, $SD = 5.16$; range = 34–52 students) than in the Netherlands ($M_{\text{classroom size}} = 23$ students, $SD = 3.68$; range = 8–29 students; $t(1009.25) = -77.30$, $p < .001$). Therefore, we controlled for Age and Classroom Size in the analyses.

5.2 Procedure

Approval for the Dutch data collection was obtained from the Ethics Review Board of the University of (blinded for review). As China has no official Ethics Review Board, an independent senior researcher in China reviewed our research plan and confirmed that it complied with Chinese law. In both countries, students' parents received information letters and could object to their children's participation. Students filled out a questionnaire about their relationship with their teacher and their engagement with schoolwork. The total questionnaire took approximately 30 minutes to complete. Teachers were asked to leave the classroom while students completed the questionnaire to stimulate free and honest responses.

5.3 Instruments

5.3.1 Student–Teacher Relationships

Students reported about the affective quality of the relationship with their teacher on the Closeness and Conflict subscales of the Student Perception of Affective Relationship with Teacher Scale (SPARTS; Koomen & Jellesma, 2015). Example items for Closeness (eight items) are “I tell my teacher things that are important to me” and “My teacher understands me”. Example items for Conflict (ten items) are “I easily have quarrels with my teacher” and “My teacher treats me unfairly”. Items were answered on a 5–point Likert–type scale, ranging from 1 (*No, that is not true*) to 5 (*Yes, that is true*). Previous studies have supported the reliability and validity of both the Dutch and Chinese version of the SPARTS (Chen et al., 2019; Koomen & Jellesma, 2015; Jellesma et al., 2015). In the present study, Cronbach's alphas ranged from .72 to .84 (see Table 19.1).

5.3.2 Engagement with Schoolwork

Students rated their engagement with schoolwork on the Behavioral and Emotional Engagement subscales of the Engagement versus Disaffection with Learning Questionnaire (Skinner et al., 2008; Dutch translation and adaptation by Zee & Koomen, 2019). Behavioral Engagement consists of six items, such as “I try hard to do well in school” and “When I am in class, I just act like I'm working” (reverse coded). Emotional Engagement includes five items, such as “I enjoy learning new

Table 19.1 Means (M), standard deviations (SD), internal consistencies (α) and correlations between main variables per sample

	Dutch sample			Chinese sample			1.	2.	3.	4.
	<i>M</i> (<i>SD</i>)	range	α	<i>M</i> (<i>SD</i>)	range	α				
1. Closeness	3.50 (0.86)	1.00– 5.00	.84	4.04 (0.84)	1.13– 5.00	.84	–	–.45**	.57**	.64**
2. Conflict	1.70 (0.67)	1.00– 4.50	.83	1.55 (0.52)	1.00– 5.00	.72	–.56**	–	–.45**	–.52**
3. Behavioral engagement	4.24 (0.58)	1.33– 5.00	.76	4.26 (0.73)	1.33– 5.00	.81	.37**	–.43**	–	.69**
4. Emotional engagement	3.79 (0.75)	1.40– 5.00	.62	4.40 (0.76)	1.00– 5.00	.80	.49**	–.48**	.57**	–

Note. * $p < .05$. ** $p < .01$. Descriptives and correlations for the Dutch sample are below the diagonal; descriptives and correlations for the Chinese sample are above the diagonal

things in class” and “When we work on something in class, I feel bored” (reverse coded). Students answered the items on a 5–point scale, varying from 1 (*No, that is not true*) to 5 (*Yes, that is true*). Items were translated in Chinese with a back translation procedure. The back translation procedure indicated that the formulation of two items needed to be slightly adapted to correspondent sufficiently with the original items, which are in English: “When I am in class, I listen very carefully” and “In class, I work as hard as I can”.

Support has been found for the reliability and validity of the Engagement Questionnaire in Western contexts (Skinner et al., 2008; Zee & Koomen, 2019). In the present study, we found evidence for partial scalar measurement invariance across the Dutch and Chinese samples ($\chi^2(96) = 298.877, p < .001$; RMSEA = .055; CFI = .915; SRMR = .069). Partial scalar invariance is considered to be sufficient to make meaningful cross–cultural comparisons (Little, 2013). In the present sample, internal consistencies varied from .62 to .81 (see Table 19.1).

5.4 Analyses

Data were analyzed in SPSS Statistics version 25. As students were nested within classrooms, we used hierarchical linear modeling with two levels (i.e., student level and classroom level) to analyze the data. We built separate models for Behavioral Engagement and Emotional Engagement. In both models, Closeness, Conflict, Sample (0 = Dutch sample, 1 = Chinese sample), and the interaction effects between Closeness and Sample and between Conflict and Sample were included as independent variables. The two interaction effects were included to investigate whether the strength of associations between student–teacher relationships and engagement differed across samples. Classroom Size, Age (in years), and students’ Gender (0 = boys, 1 = girls) were included as covariates in the analyses. To ease interpretation of results, all continuous variables were standardized at the student level (z -scores).

6 Results

Table 19.1 provides the descriptive statistics and correlations between the main study variables per sample. In both samples, the correlations between Closeness and both Behavioral and Emotional Engagement were significant and positive ($r_s = .37-.64, p_s < .05$), whereas the correlations between Conflict and the two Engagement dimensions were significant and negative ($r_s = -.43 - -.52, p_s < .05$).

In Table 19.2, the multilevel associations between the affective quality of student–teacher relationships and students’ engagement can be found. Closeness was positively associated with Behavioral Engagement ($\beta = .18, p < .001$) and Emotional Engagement ($\beta = .33, p < .001$). Furthermore, significant interaction effects between Closeness and Sample were found for both Engagement dimensions ($\beta = .36, p < .001$ and $\beta = .17, p = .001$, respectively). Figure 19.1a shows that the association between Closeness and Behavioral Engagement was stronger in the Chinese sample

Table 19.2 Associations between student–teacher relationships and students’ engagement

	Behavioral engagement	Emotional engagement
	β (SE)	β (SE)
Classroom size	.07 (.08)	.05 (.08)
Gender (0 = boy; 1 = girl)	.15 (.05)**	.05 (.04)
Age	-.01 (.03)	.06 (.03)
Closeness	.18 (.04)**	.33 (.03)**
Conflict	-.27 (.03)**	-.24 (.03)**
Sample (0 = Dutch; 1 = Chinese)	-.40 (.17)*	.28 (.16)
Closeness*sample	.36 (.06)**	.17 (.05)**
Conflict*sample	-.06 (.06)	-.09 (.05)
<i>Variance</i>		
Students	.66	.50
Classrooms	.03	.03

Notes. Standardized regression coefficients are reported. * $p < .05$. ** $p < .01$

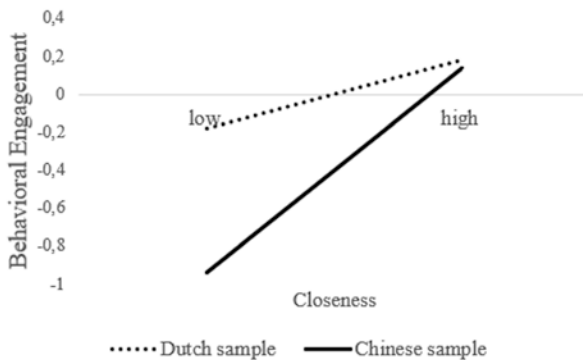


Fig. 19.1a Interaction effect of closeness and sample on behavioral engagement

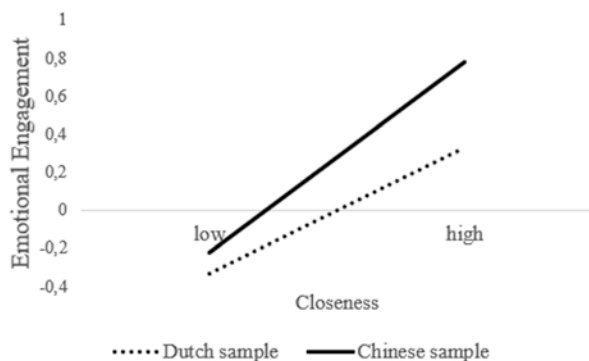


Fig. 19.1b Interaction effect of closeness and sample on emotional engagement

than in the Dutch sample. Figure 19.1b reveals that the association between Closeness and Emotional Engagement was also stronger in the Chinese sample. Conflict was negatively associated with both Behavioral Engagement ($\beta = -.27$, $p < .001$) and Emotional Engagement ($\beta = -.24$, $p < .001$). The interaction effects between Conflict and Sample were not significant for both Engagement dimensions ($\beta = -.06$, $p = .318$ and $\beta = -.09$, $p = .075$, respectively), indicating that the associations between Conflict and both Behavioral and Emotional Engagement did not differ across samples.

7 Discussion and Conclusion

In the present study, we compared students from China (an Eastern, collectivistic country) and the Netherlands (a Western, individualistic country). Specifically, we examined the extent to which associations between the affective quality of dyadic student–teacher relationships and students’ engagement differed between the two countries.

7.1 Cross–Cultural Differences in Associations

As expected (Hofstede et al., 2010; Triandis, 2001, 2018; Zhou et al., 2012), associations between student–teacher closeness and students’ engagement were stronger in the Chinese sample than in the Dutch sample. This cultural difference in strength of associations was found for both students’ engaged behaviors (cf., Zhou et al., 2012) as well as their engaged emotions (cf., Lei et al., 2018, for negative academic emotions), providing relatively strong evidence for this finding. As such, our findings support the idea that the degree of warmth, trust, and open

communication in students' relationships with their teachers is more influential for the behavioral and emotional engagement of Chinese students, most likely, because of the higher levels of interpersonal interdependency in the Chinese society compared to Dutch society (Hofstede et al., 2010; Triandis, 2001, 2018).

In contrast, associations between student–teacher conflict and students' engagement were just as strong in the Chinese sample as in the Dutch sample. Again, this was true for both the degree of effort, persistence, and concentration students put into their schoolwork (behavioral engagement) and for the feelings and emotions they experienced while working on their schoolwork (emotional engagement). Despite the potentially larger power distance and more respect for authority in Chinese schools and the broader society (Hofstede et al., 2010; Hofstede Insights, n.d.; Li, 2010), conflict did not appear to be more influential for students' engagement than in Dutch schools with a smaller power distance and less respect for authority. A possible explanation could be that high levels of negativity, tension, and hostility in relationships with teachers is harmful in all countries regardless of the specific cultural values in schools and the broader society (cf., Roorda et al., 2011; Ryan & Deci, 2017). Supporting this idea, studies conducted in Western countries usually find that conflict is more strongly associated with multiple aspects of elementary students' school adjustment (e.g., engagement, achievement, externalizing behavior; Hamre & Pianta, 2001; Lei et al., 2016; Roorda et al., 2011) than closeness. Hence, it might be that the negative impact of student–teacher conflict is more universal, whereas the impact of student–teacher closeness depends more on the cultural values and opinions existent in the specific school context and the society as a whole. More cross-cultural research, including other countries as well, is needed to further investigate this hypothesis.

7.2 *Limitations*

Some limitations need to be considered when interpreting the findings of the present study. First, we used a cross-sectional design, which does not permit statements about causality of influences. Our decision to view the student–teacher relationship as independent variable was based on both leading theories and existing research (Roorda et al., 2017; Verschueren & Koomen, 2012). Still, some studies suggest that students' engagement with schoolwork may impact the relationships they develop with their teachers as well (e.g., Zee et al., 2020). Cross-cultural studies with a longitudinal design are needed to examine the direction of influences and whether these differ across countries.

Second, students reported about both student–teacher relationship quality and their engagement with schoolwork. As most studies in elementary school are based on teachers' relationship perceptions (Roorda et al., 2011) and students tend to have different relationship perceptions than teachers (Hughes, 2011; Koomen & Jellesma, 2015), our focus on students' relationship perceptions can be considered as a strong point. Still, associations might be overestimated due to same-informant bias

(Roorda et al., 2011). Cross-cultural studies including both teacher-reports and student-reports about relationship quality and students’ engagement would therefore be helpful.

Third, our study focused on upper elementary students and only included students from China and the Netherlands. More cross-cultural research, including younger and older students and students (and teachers) from other countries is needed to find out whether our results can be generalized to different school grades and countries.

7.3 Implications for Research and School Practice

Despite these limitations, our study has several implications for future research. First, our study is a further confirmation of the idea that associations between student–teacher relationships and students’ school adjustment differ across cultures (cf., Lei et al., 2018; Zhou et al., 2012). Other cross-cultural studies focusing on dyadic student–teacher relationships also found different results for Eastern, collectivistic samples compared to Western, individualistic samples. For instance, students and teachers from Eastern, collectivistic countries appear to experience more closeness and less conflict in their mutual relationships than their Western counterparts (e.g., Beyazkurk & Kesner, 2005; Chen et al., 2019; Jia et al., 2009; Yang et al., 2013). Taken together, these studies suggest that findings from Western, individualistic contexts cannot simply be generalized to Eastern, collectivistic contexts. More research on student–teacher relationships in Eastern, collectivistic countries as well as cross-cultural comparison studies are therefore needed.

Second, previous studies found evidence for cross-cultural differences in associations between positive relationship dimensions and students’ engagement (Lei et al., 2018; Zhou et al., 2012) but did not look into negative relationship dimensions (e.g., conflict). Our present findings, however, seem to imply that there are cultural differences in the importance of positive dimensions (closeness) for students’ engagement but that the importance of conflict might be comparable across cultures. For future cross-cultural studies, it therefore seems to be important to include negative relationship dimensions, such as student–teacher conflict, as well.

The present study also has some implications for teachers and school practitioners. First, conflict appeared to be associated with both students’ behavioral and emotional engagement and these associations were just as strong in China as in the Netherlands. For both countries, it thus seems to be equally important to make teachers and other school practitioners aware of the negative impact that conflict can have on their students’ engagement with schoolwork and, hence, on their academic achievement (Roorda et al., 2017). To prevent these negative influences, teachers would profit from professional help to improve highly conflictual relationships with their students. For the Dutch context, a short teacher-based coaching intervention is available, called Teacher Student Interaction Coaching (LLInC; Bosman et al., 2021; Spilt et al., 2012). This intervention has been found effective in diminishing

conflict and increasing closeness between Dutch teachers and students (Bosman et al., 2021; Spilt et al., 2012). More research is needed, however, to investigate whether LLInC and other Western interventions (see Kincade et al., 2020, for a meta-analysis), will also be effective in Eastern, collectivistic countries. Cultural differences in prevailing expectations and norms for teacher and student behaviors (Hofstede et al., 2010) and student–teacher relationship quality (Beyazkurk & Kesner, 2005; Chen et al., 2019; Jia et al., 2009; Yang et al., 2013) suggest that Western interventions may not be automatically applicable in Eastern school contexts.

Second, associations between student–teacher closeness and students’ engagement appeared to be stronger in China than in the Netherlands. For Chinese teachers, it therefore seems to be even more important to invest in developing close and warm relationships with students than for their Dutch counterparts. If students and teachers do not succeed in developing warm, close relationships with each other, intervention programs might help. As far as we know, intervention programs focusing on increasing closeness in dyadic student–teacher relationships do not yet exist for the Chinese school context. Therefore, existing, Western programs might be adapted for the Chinese context (Bosman et al., 2021; Kincade et al., 2020) or new interventions might be developed especially designed for Chinese schools. For Dutch teachers, this finding may also have implications. More specifically, it might be that student–teacher closeness is also more important for the engagement of students with a Chinese background in Dutch schools and, hence, investing in warm, close relationships may also be more important for these students. More research is needed, however, to find out whether our findings generalize to Chinese students in Western school contexts as well. In addition, future cross-cultural studies, including other countries and using longitudinal designs, could provide more insight in cultural differences in the associations between the affective quality of dyadic student–teacher relationships and students’ engagement with schoolwork.

References

- Archambault, I., Pagani, L. S., & Fitzpatrick, C. (2013). Transactional associations between classroom engagement and relations with teachers from first through fourth grade. *Learning and Instruction, 23*, 1–9. <https://doi.org/10.1016/j.learninstruc.2012.09.003>
- Beyazkurk, D., & Kesner, J. E. (2005). Teacher-child relationship in Turkish and United States school: A cross-cultural study. *International Education Journal, 6*(5), 574–554. <http://iej.cjb.net>
- Bosman, R. J., Zee, M., de Jong, P. F., & Koomen, H. M. Y. (2021). Using relationship-focused reflection to improve teacher-child relationships and teachers’ student-specific self-efficacy. *Journal of School Psychology, 87*, 28–47. <https://doi.org/10.1016/j.jsp.2021.06.001>
- Chen, M., Zee, M., Koomen, H. M. Y., & Roorda, D. L. (2019). Understanding cross-cultural differences in affective teacher–student relationships: A comparison between Dutch and Chinese primary school teachers and students. *Journal of School Psychology, 76*, 89–106. <https://doi.org/10.1016/j.jsp.2019.07.011>
- Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children’s school outcomes through eighth grade. *Child Development, 72*(2), 625–638. <https://doi.org/10.1111/1467-8624.00301>

- Hoang, N., Holopainen, L., & Siekkinen, M. (2018). Quality of teacher-child interactions and its relations to children’s classroom engagement and disaffection in Vietnamese kindergartners. *International Journal of Early Years Education*, 26(4), 387–402. <https://doi.org/10.1080/009669760.2018.1478281>
- Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: Software of the mind* (Rev. 3rd ed.). McGraw–Hill.
- Hofstede Insights. (n.d.). Country Comparison. <https://hofstede-insights.com/country-comparison/>
- Hu, B. Y., Teo, T., Nie, Y., & Wu, Z. (2017). Classroom quality and Chinese preschool children’s approaches to learning. *Learning and Individual Differences*, 54, 51–59. <https://doi.org/10.1016/j.lindif.2017.01.007>
- Hu, B. Y., Wu, Z., Winsler, A., Wu, Y., & Fan, X. (2021). Teacher-child interaction and preschoolers’ learning behavior in China: A piecewise growth model. *Early Education and Development*, 32(5), 677–694. <https://doi.org/10.1080/10409289.2020.1798719>
- Hughes, J. N. (2011). Longitudinal effects of teacher and student perceptions of teacher–student relationship qualities on academic adjustment. *The Elementary School Journal*, 112(1), 38–60. <https://doi.org/10.1086/660686>
- Jellesma, F. C., Zee, M., & Koomen, H. M. Y. (2015). Children’s perceptions of the relationship with the teacher: Associations with appraisals and internalizing problems in middle childhood. *Journal of Applied Developmental Psychology*, 36, 30–38. <https://doi.org/10.1016/j.appdev.2014.09.002>
- Jia, Y., Ling, G., Chen, X., Ke, X., Way, N., Hughes, D., & Lu, Z. (2009). The influence of student perceptions of school climate on socioemotional and academic adjustment: A comparison of Chinese and American adolescents. *Child Development*, 80(5), 1514–1530. <https://doi.org/10.1111/j.1467-8624.2009.01348.x>
- Kincade, L., Cook, C., & Goerd, A. (2020). Meta-analysis and common practice elements of universal approaches to improving student-teacher relationships. *Review of Educational Research*, 90(5), 710–748. <https://doi.org/10.3102/0034654320946836>
- Koomen, H. M. Y., & Jellesma, F. C. (2015). Can closeness, conflict, and dependency be used to characterize students’ perceptions of the affective relationship with their teacher? Testing a new child measure in middle childhood. *British Journal of Educational Psychology*, 85(4), 479–497. <https://doi.org/10.1111/bjep.12094>
- Lei, H., Cui, Y., & Chiu, M. M. (2016). Affective teacher-student relationships and students’ externalizing behavior problems: A meta-analysis. *Frontiers in Psychology*, 7, 1–12. <https://doi.org/10.3389/fpsyg.2016.01311>
- Lei, H., Cui, Y., & Chiu, M. M. (2018). The relationship between teacher support and students’ academic emotions: A meta-analysis. *Frontiers in Psychology*, 8, 1–12. <https://doi.org/10.3389/fpsyg.2017.02288>
- Li, J. (2010). Learning to self-perfect: Chinese beliefs about learning. In C. Chan & N. Rao (Eds.), *Revisiting the Chinese learner. CERC studies in comparative education* (Vol. 25, pp. 35–69). https://doi.org/10.1007/978-90-481-3840-1_2
- Little, T. D. (2013). *Longitudinal structural equation modeling*. Guilford Press.
- Pianta, R. C. (1999). *Enhancing relationships between children and teachers*. American Psychological Association.
- Pianta, R. C., Hamre, B., & Stuhlman, M. (2003). Relationships between teachers and children. In W. M. Reynolds, G. E. Miller, & I. B. Weiner (Eds.), *Handbook of psychology: Educational psychology* (Vol. 7, pp. 199–234). John Wiley & Sons.
- Roorda, D. L., Jak, S., Zee, M., Oort, F. J., & Koomen, H. M. Y. (2017). Affective teacher–student relationships and students’ engagement and achievement: A meta-analytic update and test of the mediating role of engagement. *School Psychology Review*, 46(3), 1–23. <https://doi.org/10.17105/SPR-2017-0035.V46-3>
- Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher–student relationships on students’ school engagement and achievement: A meta-analytic approach. *Review of Educational Research*, 81(4), 493–529. <https://doi.org/10.3102/0034654311421793>

- Ryan, R. M., & Deci, E. L. (2017). Pervasive social influences part I: Cultural contexts. In R. M. Ryan & E. L. Deci (Eds.), *Self-determination theory: Basic psychological needs in motivation, development, and wellness* (pp. 561–590). The Guilford Press.
- Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology, 100*(4), 765–781. <https://doi.org/10.1037/a0012840>
- Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement, 69*(3), 493–525. <https://doi.org/10.1177/0013164408323233>
- Spilt, J. L., Koomen, H. M. Y., Thijs, J. T., & van der Leij, A. (2012). Supporting teachers' relationship with disruptive children: The potential of relationship focused intervention. *Attachment & Human Development, 14*(3), 305–318. <https://doi.org/10.1080/14616734.2012.672286>
- Triandis, H. C. (2001). Individualism–collectivism and personality. *Journal of Personality, 69*(6), 907–924. <https://doi.org/10.1111/1467-6494.696169>
- Triandis, H. C. (2018). *Individualism and collectivism*. Routledge. <https://doi.org/10.4324/9780429499845>
- Triandis, H. C., Bontempo, R., Villareal, M. J., & Asai, M. (1988). Individualism and collectivism: Cross–Cultural perspectives on self–group relationships. *Journal of Personality and Social Psychology, 54*(2), 323–338. <https://doi.org/10.1037/0022-3514.54.2.323>
- Verschueren, K., & Koomen, H. M. Y. (2012). Teacher–Child relationships from an attachment perspective. *Attachment & Human Development, 14*(3), 205–211. <https://doi.org/10.1080/14616734.2012.672260>
- Yang, C., Bear, G. G., Chen, F. F., Zhang, W., Blank, J. C., & Huang, X. (2013). Students' perceptions of school climate in the U.S. and China. *School Psychology Quarterly, 28*(1), 7–24. <https://doi.org/10.1037/spq0000002>
- Zee, M., & Koomen, H. M. Y. (2019). Engaging children in the upper elementary grades: Unique contributions of teacher self–efficacy, autonomy support, and student–teacher relationships. *Journal of Research in Childhood Education, 34*(4), 477–495. <https://doi.org/10.1080/02568543.2019.1701589>
- Zee, M., Rudasill, K. M., & Bosman, R. J. (2020). A cross-lagged study of students' motivation, academic achievement, and relationships with teachers from kindergarten to 6th grade. *Journal of Educational Psychology, 113*, 1208–1226. Advance online publication. doi:<https://doi.org/10.1037/edu0000574>.
- Zhou, N., Lam, S.-F., & Chan, K. C. (2012). The Chinese classroom paradox: A cross–cultural comparison of teacher controlling behaviors. *Journal of Educational Psychology, 104*(4), 1162–1174. <https://doi.org/10.1037/a0027609>

Debora Roorda is assistant professor at the Research Institute of Child Development and Education, University of Amsterdam. Her research is inspired by both attachment and interpersonal theory and focuses on student–teacher relationships in primary and secondary school and their impact on students' academic and behavioral adjustment and teachers' wellbeing. She also examines, amongst others, how student characteristics (gender, problem behaviors) influence student–teacher relationship quality and whether the complementarity principle applies to student–teacher interactions. Furthermore, she is interested in instrument development (e.g., relationship drawings, priming studies) and intervention effectiveness. Debora publishes regularly in various high-impact journals.

Mengdi Chen is a research assistant professor at the Faculty of Education, University of Macau. Her research interest lies on cultural differences in affective student-teacher relationship quality, the associations between student characteristics (e.g., gender, age, shyness) and student-teacher relationships, and the application of teacher-student relationship drawings in cross-cultural contexts.

Marjolein Zee is an associate professor at the Department of Psychology, Education and Child Studies, Erasmus University Rotterdam, Rotterdam, The Netherlands. Combining a variety of theories with advanced statistical methods, she aims to unravel how teachers’ biases, (self-efficacy) beliefs, feelings, and mental representations of relationships affect their behaviors toward individual students with various behavioral problems and needs, including students with self-regulation problems. Zee has developed several instruments and coding schemes (student-specific teacher self-efficacy measure, relationship drawings) and her research results have been published in numerous peer-reviewed and popular articles.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.


The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 20

The Mediated Relationship Between Secondary School Student Perceptions of Teaching Behaviour and Self-Reported Academic Engagement Across Six Countries



Ridwan Maulana , Rikkert van der Lans, Michelle Helms-Lorenz , Sibel Telli, Yulia Irnidayanti, Nurul Fadhilah, Carmen-Maria Fernandez-Garcia, Mercedes Inda-Caro, Seyeoung Chun, Okhwa Lee, Thelma de Jager, and Thys Coetzee

Abstract Limitations in the current knowledge base on the importance of perceived teaching behaviour and student engagement are visible. Past studies on this topic specifically take place in certain contexts (usually the Western context) using various instruments. The current study aims to extend our understanding of the link between perceived teaching behaviour and student engagement based on students' perceptions using uniform measures across six contrasting national contexts. It also aims to explore the role of certain personal variables in the interplay between students' perceived teaching behaviour and engagement. In total, 40,788 students in The Netherlands, Spain, Indonesia, South Korea, South Africa, and Turkey participated in the survey using the My Teacher Questionnaire (MTQ) and the Student Engagement scale. Item Response Theory (IRT) and Classical Test Theory (CTT)

R. Maulana (✉) · M. Helms-Lorenz
Department of Teacher Education, University of Groningen, Groningen, The Netherlands
e-mail: r.maulana@rug.nl; m.helms-lorenz@rug.nl

R. van der Lans
Curium – Leiden University Medical Centre, Leiden, Netherlands
University of Amsterdam, Amsterdam, Netherlands

S. Telli
Çanakkale Onsekiz Mart University, Çanakkale, Turkey

Y. Irnidayanti
Department of Biology and Biology Education, Faculty of Mathematics and Science, State University of Jakarta, Jakarta, Indonesia

N. Fadhilah
Department of Biostatistics and Population Studies, Faculty of Public Health, Universitas Indonesia, Depok, Indonesia

analyses were used to analyse the student data. Results show that, in general, perceived teaching behaviour is positively related, and mostly strongly, to student engagement across the six educational contexts. This means the higher the perceived teaching behaviour, the higher students reported their academic engagement, and vice versa. Slight differences in the magnitude of relationships between perceived teaching behaviour and engagement are evident. The strongest link was found in the Netherlands, followed by South Korea, South Africa, Indonesia, Turkey, and Spain. Student gender, age, and school subject hardly show effects on the interplay between perceived teaching behaviour and engagement. Implications for research and practice are discussed.

Keywords Cross-national study · Secondary education · Student engagement · Student perceptions · Teaching behaviour

1 Introduction

Effective teaching behaviour is of key importance for student learning and outcomes (Coe et al., 2014; Hattie, 2009; Muijs et al., 2014). Among other student learning characteristics, student engagement has been recognized as an important predictor of students' academic performance (Appleton et al., 2008). Student engagement is viewed as the primary theoretical model for promoting school completion characterized by sufficient academic and social skills to contribute in concurrent and subsequent academic success (Christenson et al., 2008; Finn, 2006; Skinner et al., 2008). Furthermore, student engagement has been identified as a powerful mediator between teaching quality and student achievement (Virtanen et al., 2015). Research shows that student engagement is positively associated with various aspects of effective teaching within one educational context (Maulana et al., 2017; Pianta et al., 2012; Rimm-Kaufman et al., 2015). Higher levels of

C.-M. Fernandez-Garcia

Department of Educational Sciences, University of Oviedo, Oviedo, Spain

e-mail: fernandezcarmen@uniovi.es

M. Inda-Caro

University of Oviedo, Oviedo, Spain

e-mail: indamaria@uniovi.es

S. Chun

Department of Education, Chungnam National University, Daejeon, South Korea

O. Lee

Department of Education, Chungbuk National University, Cheongju, South Korea

T. de Jager · T. Coetzee

The Department of Educational Foundation, Tshwane University of Technology,

Pretoria, South Africa

engagement are uniquely associated with higher levels of teaching quality (Quin et al., 2017; Wang & Eccles, 2012). Because teacher-student interaction is recognized as a primary source of student development (Pianta & Allen, 2008), understanding the universal link between teaching behaviour and student engagement is a desideratum. This is because social interactions with teachers within the school setting may serve as a protective factor for students who are weakly engaged in learning (Guo et al., 2011).

Despite evidence on the importance of teachers' teaching quality for student engagement, the current knowledge base is limited in at least three ways. First, most studies on engagement and teaching behaviour have been conducted in the West. Hence, the extent to which the results of Western studies represent non-Western contexts is unclear. Particularly, little is known whether the link between perceived teaching behaviour and student engagement differs in magnitude across different educational contexts. Next, it remains unclear whether certain personal (i.e., student gender and age) and contextual (i.e., school subject) factors can explain the differential link between perceived teaching behaviour and engagement in Western and non-Western contexts. Some studies indicate direct links between student gender, age, and school subjects on perceived teaching behaviour or on self-reported engagement separately (e.g., Cohen et al., 2018; Cooper, 2014; Fernández-García et al., 2019; Havik & Westergard, 2020; Lietaert et al., 2015). Identifying potential mediating roles of student gender, age, and school subject in the relationship between perceived teaching behaviour and engagement is important to inform more tailored interventions for teaching improvement.

Previous studies on teaching behaviour and student engagement are rather fragmented and are mostly restricted to a single context (e.g., Maulana et al., 2017; Roorda et al., 2011; Virtanen et al., 2015). Although there were studies conducted in non-Western contexts, these were typically published in their own languages and published in local, non-English, journals (e.g., Hidayati & Rodliyah, 2020). Hence, it is largely unknown whether the link between perceived engagement and teaching behaviour, and the role of some background variables, is universal.

Furthermore, past studies typically studied teaching behaviour and student engagement using various instruments. The instruments used in past studies vary, at least to certain degree, concerning their underlying conceptualizations, operationalisations, and modes (observation vs. self-reports). The heterogeneity of the instruments poses challenges for comparing the link between teaching behaviour across contexts more accurately. In addition, research examining the association between student perceptions of teaching behaviour and their perceived engagement across various contexts is scarce (Quin et al., 2017).

The present study aims to empirically extend our understanding of the link between perceived teaching behaviour and student engagement based on students' perceptions using uniform measures across six contrasting national contexts. In addition, it also aims to explore the role of certain personal variables in the interplay between perceived teaching behaviour and engagement cross-nationally. The study includes representatives of both Western and Eastern contexts.

2 Literature Review

2.1 Teaching Behaviour

This study applies a conceptualization of teaching behaviour that is grounded in the teaching and teacher effectiveness literature (e.g., Hattie, 2009; Muijs et al., 2014; Van de Grift, 2014). Teaching behaviour refers to teachers' acts contributing to student learning and outcomes (Maulana et al., 2021). Some examples include showing respects in the learning process, providing students with clear examples, and requesting students to reflect on their learning approaches.

Typically, the variety in effective teaching behaviours is grouped into seven broader domains (Bell et al., 2019; Muijs et al., 2014). Past research in Indonesia, South Korea, the Netherlands, South Africa, Spain, and Turkey shows that the variety in student perceptions of effective teaching behaviors can be represented by a six-factor structure labelled as: *safe and stimulating learning climate, efficient classroom management, clarity of instruction, activating teaching, teaching learning strategies, and differentiation* (André et al., 2020; Inda-Caro et al., 2019; Maulana & Helms-Lorenz, 2016).

Scholars show that teaching behaviours, based on observation data in the Netherlands, can be ordered hierarchically along a latent continuum (Van de Grift et al., 2011). This complementary conceptualization of teaching behaviour is grounded in theories on teacher development proposed by Berliner (2004) and Fuller (1969). Other scholars show that the unidimensionality of teaching behaviour, based on student perceptions data in the Netherlands, is confirmed (Maulana et al., 2015a; Van der Lans & Maulana, 2018; Van der Lans et al., 2015). The current evidence-base has been extended to other cultural contexts including Indonesia, South Korea, South Africa, Spain, and Turkey (Maulana et al., 2015b; Van der Lans et al., 2021).

2.2 Student Engagement

Student engagement is multifaceted and multidimensional in nature (Alrashidi et al., 2016). It is frequently conceptualized as the extent to which students are behaviourally and psychologically engaged in academic tasks (Appleton et al., 2006; Van de Grift, 2007; Wang & Holcombe, 2010). *Behavioural* engagement refers to participation and involvement in academic, social or extracurricular activities, shown in terms of attendance, time spent on assignments, concentration, and attention. The focus is on students' actions and practices that are directed toward school and learning (e.g., The student tries to work hard in class, shows a positive conduct and effort, participates in class discussions, follows the rules, and pays attention). Behavioural engagement is important for achieving positive academic outcomes and for preventing school dropout.

Emotional engagement refers to the extent of positive and negative reactions to teachers, classmates, academics and school, shown in terms of interest and positive attitude. The emphasis is on students' affective reactions and sense of identification with school (e.g., how students feel in the classroom, whether they enjoy learning new things, get involved when they are working on something or show interest (Fredricks et al., 2004; Jimerson et al., 2003; Wang & Holcombe, 2010). Emotional engagement is linked to strengthening ties to school and willingness to perform academic work.

2.3 Perceived Teaching Behaviour and Engagement

Student engagement is a malleable variable that depends on environmental contexts (Marks, 2000; Pianta et al., 2012). Instead of solely associating it with a trait of an individual, the contemporary definition of engagement emphasizes the interaction between an individual's characteristics and their environments (Thijs & Verkuyten, 2009). Previous studies have shown that teaching behaviour contributes to a range of student outcomes (e.g., Pianta & Allen, 2008). Teachers' provision of emotional support to students has been linked with various positive academic outcomes including social skills and academic competence (Malecki & Demaray, 2003), teacher reports of high levels of student participation in class, students' self-reports of engagement and task completion (Anderson et al., 2004), subjective well-being (Suldo et al., 2009), school satisfaction (Richman et al., 1998), experience of meaningfulness of schoolwork and on-task orientation (Thuen & Bru, 2000).

Studies show that student perceptions of teaching behaviour has a powerful effect on students' self-report of cognitive and behavioural engagement (Bertills et al., 2019; Davidson et al., 2010). Particularly, Inda-Caro et al. (2019) found that perceived emotional engagement was more strongly related to perceived teaching behaviour than perceived behavioural engagement. Furthermore, Klem and Connell (2004) found that perceived teaching behaviour was also related to classroom and school engagement (effort and attention in classes, being prepared for classes and finding school personally important). Ryan and Patrick (2001) found that perceived classroom emotional support was related positively to perceived engagement in self-regulated learning, and negatively to off-task and disruptive behaviour in the classroom. Den Brok et al. (2005) show that perceived teacher friendliness in the classroom is associated with perceived willingness to put effort into learning the school subject.

Taken together, there is evidence that student perceptions of effective teaching behaviour contribute positively to their perceptions of various academic outcomes. In classrooms with highly effective teachers, students' needs of relatedness, competence, and autonomy are met (Deci & Ryan, 2000), which is reflected in student behavioural and emotional engagement and successful learning (Virtanen et al., 2015). However, there is evidence that teaching behaviour in secondary schools varies between classrooms (Malmberg et al., 2010; Maulana et al., 2015b) and between

countries (Maulana et al., 2021), suggesting that not all secondary school students perceive high-quality teaching and learning at all times. Therefore, this variation is to be expected at the classroom level across countries.

2.4 Perceptions of Teaching Behaviour and Student Engagement: Gender, Age, and School Subject

The present study focuses on secondary school students. These students are in the adolescent period, which is characterized by changes in biological, cognitive, emotional, and social reorganisation (Susman & Rogol, 2013). Hence, their personal characteristics may play a role in the interplay between perceived teaching behaviour and engagement. A limited number of studies investigate the link between certain personal characteristics with either engagement or teaching behaviour separately. To our knowledge, studies examining the mediating effect of students' characteristics on the two key constructs are underrepresented in the literature. The current study aims to test the mediating effect of several personal background on the relationship between perceived teaching quality and engagement across countries. Thus, student gender, age, and school subject are focused on.

2.4.1 Student Gender and Engagement

In general, research has consistently shown that boys show lower academic engagement than girls. This trend is consistent across primary, secondary, and higher education. For example, Cooper (2014) found that in Grades 9–12, boys reported lower engagement than girls in the United States. A large-scale study involving 7th–9th graders in 12 countries (Austria, Canada, China, Cyprus, Estonia, Greece, Malta, Portugal, Romania, South Korea, the United Kingdom, and the United States) show a similar trend (Lam et al., 2014). Jelas et al. (2014) and Amir et al. (2014) confirmed similar finding among 12–16 years old students in Malaysia. This trend was also visible among university students in Malaysia (Teoh et al., 2013). More recent studies involving primary school students in Japan (Oga-Baldwin & Fryer, 2020), primary and lower secondary school in Norway (Havik & Westergard, 2020), upper secondary school students in China (Teuber et al., 2021), and a wide range of age-groups (12–25 years old) in Portugal (Santos et al., 2021) show consistent findings.

In secondary education, academic engagement tends to decline for both boys and girls over time. This trend was found in Flanders (Dutch-speaking part of Belgium) (Lamote et al., 2013; Van de Gaer et al., 2009), The Netherlands (Opdenakker et al., 2012), and the United States (Wang & Eccles, 2012). A larger decline in engagement was found for boys than for girls in Canada (Chouinard & Roy, 2008), The United States (Dotterer et al., 2009), and Australia (Watt, 2000). A meta-analysis study shows that, under equal levels of intellectual ability, girls are more likely to be

academically successful because they engage more in schoolwork than boys over time (Lei et al., 2018).

Existing studies on gender and engagement are typically fragmented including single contexts. An exception is the study of Lam et al. (2012) studying gender differences in engagement among 7th and 9th graders across 12 countries. Lam et al. found that girls reported higher levels of school engagement compared to boys. The teachers rated girls higher in academic performance compared to boys. However, student gender did not moderate the relation among student engagement, academic performance, or contextual supports (Lam et al., 2012).

2.4.2 Student Age and Engagement

Research has consistently shown that, in general, younger students tend to show higher engagement than older students. For example, younger students in primary schools reported higher emotional engagement than older students in lower secondary schools (Havik & Westergard, 2020). In a study involving students aged 12–16 years old in Malaysia, younger students reported higher engagement compared to older students (Amir et al., 2014). A similar trend was also found in Portugal involving samples of students aged 12–25 years old (Santos et al., 2021), in Canada involving secondary school students (grade 9–11) (Chouinard & Roy, 2008), and in The United States involving junior high and high school students (Dotterer et al., 2009). Students in senior grades are less likely to be interested in learning than students in junior grades (Lam et al., 2007). In a study involving 12 countries, Lam et al. (2016) found a declining trend in perceived engagement among grade 7–9 students, suggesting that perceived engagement becomes lower as students get older.

There are studies regarding the link between student gender and age and academic engagement. In general, at all grade levels in primary and secondary schools, boys consistently reveal less academic engagement than girls (Finn, 1989; Finn & Cox, 1992; Lee & Smith, 1993). More recent studies confirmed that younger female students tend to report higher levels of engagement and satisfaction with school than older male students (Amir et al., 2014; Hartono et al., 2020; Inchley et al., 2020).

2.4.3 School Subject and Engagement

Research documenting differences in engagement across school subjects are scarce. Nevertheless, it is much discussed in school practice that engaging students academically is a challenge for many teachers, irrespective of the school subject they teach. Scholars acknowledge that “engaging students in science and helping them develop an understanding of its ideas has been a consistent challenge for both science teachers and science educators alike” (Hadzigeorgiou & Schulz, 2019, p. 1). Although engagement is not a school subject-specific problem, variations in engagement across school-subjects are expected due to different characteristics of the subject (e.g., difficulty level).

2.4.4 Student Gender and Teaching Behaviour

In general, there is evidence from Western contexts that boys tend to report lower levels of teacher support (Oelsner et al., 2011; Soenens et al., 2012; Vansteenkiste et al., 2012). Girls tend to rate their teachers more favourably than boys (Furrer & Skinner, 2003). Evidence from Flemish secondary schools show that boys reported lower teacher support than girls (Lietaert et al., 2015). Similarly, Maulana et al. (2014) found that girls reported higher level of teacher support in terms of influence than did boys in Indonesian secondary schools. However, another study from Indonesia shows that student gender has no significant link with perceived teaching behaviour in terms relatedness, structure, and autonomy support (Maulana et al., 2016).

2.4.5 Student Age and Teaching Behaviour

In general, the link between student age and perceived teaching behaviour is under-represented in the literature. A study in Spain shows that students in lower secondary education rated their teachers more favourably than did their peers in upper secondary education (Fernández-García et al., 2019). A study in Indonesia shows that grade level, which corresponds to student age, has no significant link with their perceived teaching behaviour (Maulana et al., 2016).

2.4.6 School Subject and Teaching Behaviour

Given that curricular materials are largely content specific, and teacher guides for textbooks are differentially elaborated in different subjects (Remillard, 2005; Reutzel et al., 2014), differences in teaching behaviour across subjects are to be expected. The way a school subject is perceived can influence teaching (Grossman & Stodolsky, 1994). Nevertheless, the link between school subject and perceived teaching behaviour is inconclusive. In addition, there is little empirical evidence about the stability in an individual teacher's practice across different subjects (Cohen et al., 2018), particularly in secondary education. A limited number of studies in primary schools generally show that teachers' practices vary across subjects (Cohen et al., 2018; Graeber et al., 2012; Knapp et al., 1995). This suggests that the variation in the effectiveness of teaching behaviour across subjects may exist. For example, a study in the United States indicated a slightly higher quality of teaching behaviour for English teachers compared to mathematics teachers (Cohen et al., 2018). Another study examining whether students' perceptions of their teachers' interpersonal behaviour relates to students' subject-related attitudes across different school subjects from grades 9 to 11 revealed that an interpersonal Affiliation style is beneficial for all students, irrespective of the subject matter (Telli, 2016). It is important to note that the studies mentioned above used other, mostly qualitative, methods (e.g., interviews, classroom observation, teacher logs).

In secondary education, Maulana et al. (2012) found that Indonesian English as Foreign Language (EFL) teachers were perceived friendlier compared to mathematics teachers. Similarly, students in Western contexts reported that psychosocial classroom climates of math and science classes were less favourable compared to other school subjects (Den Brok et al., 2010; Levy et al., 2003). However, another study in The Netherlands reports that science and math teachers were perceived as more dominant and friendly compared to their colleagues from other school subjects (Den Brok et al., 2004). Similarly, a study in Indonesia shows that math and science teachers were perceived more positively in the provision of relatedness, structure, and autonomy support than other school subjects (Maulana et al., 2016). In addition, other scholars did not find a significant effect of school subject on perceived teaching behaviour in Indonesia (Maulana et al., 2014) and in The Netherlands (Maulana et al., 2015b).

Taken together, there seems to be a tendency that boys and older students tend to report lower engagement and less positive perceived teaching behaviour compared to girls and younger students. However, the role of school subject in explaining differences in perceived engagement and perceived teaching behaviour is inconclusive. Furthermore, studies of the mediating effect of student gender, age, and subject school on the link between perceived teaching behaviour and engagement are scarce. One study in the Flemish secondary school context examined the mediating effect of teacher support on the link between student gender and engagement, showing that teacher's provision of autonomy support and involvement partially mediated the relationship between gender and behavioural engagement (Lietaert et al., 2015). Although Lietaert et.al did not investigate the mediating role of student gender, their study suggests a potential interplay between gender, teaching behaviour, and engagement.

Based on the literature review, it is expected that student gender, student age, and school subject will play a role, at least to some degree and in certain educational contexts, in the relationship between perceived teaching behaviour and self-reported student engagement.

3 Context of the Current Study

Henrich et al. (2010) express that most of the psychological knowledge is built on studies from Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies. Most cross-cultural or cross-national studies have not directly investigated the link between teaching behaviour and engagement, and the mediating role of certain background factors, across WEIRD and non-WEIRD countries. The current study aims to examine the relationship between perceived teaching behaviour and engagement, and the mediating/moderating effect of background variables (i.e., student gender, age, and school subject) on the relationship between perceived teaching behaviour and engagement across six contrasting countries: The Netherlands, Spain, Turkey, South Africa, South Korea, and Indonesia.

This study used part of the data collected in an international project on teaching quality initiated by University of Groningen, The Netherlands. The multinational project involved 16 countries. In this study, data from six countries are included from both WEIRD (i.e., The Netherlands, Spain) non-WEIRD (South Africa, Indonesia) and in between WEIRD and non-WEIRD (South Korea, Turkey) societies, with contrasting cultural values and socio-economic development background.

3.1 Cultural Dimension

The six countries share some similarities and differences in terms of cultural dimensions and educational performance. There are at least two cultural dimensions depicting the diversity and the similarity of the six countries that are relevant to this study: Power Distance index (PDI) and Individualism versus Collectivism (IDV)¹ (Hofstede et al., 2010). Of the six countries, the Netherlands has the lowest score (PDI = 38). The Dutch society is characterized by being independent, hierarchy for convenience only, and equal rights. Superiors facilitate, empower, and are accessible. Decentralization of power is applied in which superiors count on the experience of their team members. Employees expect to be consulted. Control is disliked, attitude towards superiors is informal, and communication is direct and participative. Spain (PDI = 57), South Korea (PDI = 60), Turkey (PDI = 66) and Indonesia (PDI = 78) respectively have higher power distance scores. In high power distance countries, people are dependent on hierarchy. Superiors are directive and controlling. Centralized power is applied in which obedience to superiors is expected. Communication is indirect and people tend to avoid negative feedback (Hofstede, 2001; Hofstede et al., 2010).

Of the six countries, the Netherlands revealed the highest in IDV (80), meaning that the country is characterized by a highly individualist society. In this country, a loosely-knit social framework is highly preferred. Individuals are expected to focus on themselves and their immediate families. The superior and inferior relationship is based on mutual advantage, and meritocracy is applied as a base for hiring and promoting individuals. Management focuses on the management of individuals. The remaining countries are considered collectivistic, with Indonesia as the most collectivistic (14), followed by South Korea (18), Turkey (37), and Spain² (51) respectively. In the collectivistic society, a strongly defined social framework is highly preferred. Individuals should conform to the society's ideals and the in-groups loyalty is expected. Superior/inferior relationships are perceived in moral terms like family relationships. Management focuses on management of groups. In

¹The country data for South Africa related to these cultural values is not available. The current available data of South Africa is limited to the White population only, which is a minority group in the country.

²Spain is seen as collectivist in comparison to the rest of European societies (except Portugal), but for the rest of the world it is individualist as noted in the reference.

some collectivistic countries like Indonesia, there is a strong emphasis on (extended) family relationships, in which younger individuals are expected to respect older people and taking care of parents is highly valued (Hofstede, 2001; Hofstede et al., 2010).

With respect to educational performance, the latest worldwide study of the Programme for International Student Assessment (PISA)³ 2018 showed that South Korea's performance was well above the OECD average and listed among the top 5. The Netherlands' average performance was also above the OECD average but below the South Korean performance. Spain was positioned slightly below the OECD average. Turkey's mean performance in mathematics improved in 2018 while enrolling many more students in secondary education between 2003 and 2018 without sacrificing quality of the education. Indonesia was listed well-below the OECD average and the lowest compared to the other four countries (OECD, 2019).

4 Socio-economic Dimension

Socio-economic background is another prevalent factor in country development. Given the country's socio-economic background, variations in the link between perceived teaching behaviour and engagement, coupled with students' personal background, are expected. Thus, it is important to examine whether the direction and the magnitude of associations between perceived teaching behaviour and engagement are similar between developed (e.g., The Netherlands, South Korea, Spain, Turkey) and developing contexts (South Africa, Indonesia) (United Nations Development Programme, 2022).⁴ Past research shows that family involvement in education is more strongly related to science achievement in more developed countries (Chiu, 2007). It was also found that teacher-student relationship is more strongly associated with students' perceived classroom discipline in more developed countries (Chiu & Xihua, 2008). Studies investigating the link between perceived teaching behaviour and engagement involving developed and non-developed contexts are underrepresented. One multinational study involving 12 developed and non-developed nations indicates that most of the associations between the contextual factors (e.g., instructional practices, teacher support) and student engagement did not vary across countries (Lam et al., 2016).

Based on the literature review, it is expected that the relationship between perceived teaching behaviour and self-reported student engagement across developed and non-developed settings included in this study will not vary significantly.

³South Africa did not participate in the PISA study. Hence, the performance data for South Africa is not available.

⁴Based on Human Development Index (HDI) developed by the United Nations. HDI score of ≥ 0.80 = developed, and < 0.80 = developing. Based on the 2020 report, the HDI of the six countries is as follows respectively: The Netherlands = 0.94; South Korea = 0.92; Spain = 0.90; Turkey = 0.82; Indonesia = 0.72; South Africa = 0.71). See: <http://hdr.undp.org/en/indicators/137506>

Because it is unclear whether or not the link between perceived teaching behaviour and self-reported engagement depends on socio-economic dimension, no expectation regarding differences between developed and developing educational contexts can be made. Rather, this will be examined in an explorative manner.

5 Research Questions

This study aims to answer the following research questions:

1. How does the relationship between perceived teaching behaviour and student engagement compare between countries?
2. How does student gender, student age, and school subject mediate the relationship between student engagement and perceived teaching behaviour across countries?

6 Method

6.1 Sample

This study is part of a larger project on differentiation from an international perspective involving 16 countries, led by the University of Groningen, The Netherlands. For the present study, data from Indonesia, the Netherlands, South Africa, South Korea, Spain, and Turkey are included. In total, 40,788 secondary school students filled in the questionnaire in the six countries. Data were collected during the years 2015 (Indonesia, South Korea and the Netherlands), 2016 (South Africa and Spain), and 2017 (Turkey) using a combination of online and paper and pencil methods depending on the resources in the participating countries. Teaching behaviour and student engagement questionnaires were administered during the same period. Only the first year of data collection was included in this study. Data included perceptions of teachers with subjects from natural sciences, social sciences and languages. In all countries, data were approximately uniformly distributed across different subjects, student age and student gender (Table 20.1). Students participated in the study on a voluntary basis.

For the present study, a selection of eligible, more balance, sample was done, including Indonesia ($n = 6329$ students; 299 teachers; 24 schools), the Netherlands ($n = 6590$; 300 teachers; 148 schools), South Africa ($n = 4034$; 270 teachers; 10 schools), South Korea ($n = 6976$; 336 teachers; 26 schools), Spain ($n = 4524$; 251 teachers; 48 schools), and Turkey ($n = 7434$; 274 teachers; 16 schools). The selection comprised a random subset of five students per class for the analysis. This selection balanced the considerable variation in class size found within and between countries ($\text{Min}_{(\text{class size})} = 6$; $\text{Max}_{(\text{class size})} = 96$). Especially, the exceptionally large

Table 20.1 Distribution of school subject, student gender and student age within the six countries in the complete sample ($n_{\text{student}} = 40.788$) In brackets the distribution in the selected sample ($n_{\text{student}} = 8640$)

	Subjects			Gender	Age
	%Language	%Natural sciences	%Social sciences	(%girls)	(years)
Indonesia	21.5 (21.4)	42.0 (40.8)	36.5 (37.8)	60.0 (61.5)	16.53 (16.59)
South Korea	46.4 (46.0)	36.0 (35.6)	17.6 (18.5)	57.8 (57.7)	16.39 (16.44)
Netherlands	44.6 (43.4)	29.8 (30.2)	24.1 (24.7)	52.7 (55.0)	13.86 (13.96)
South Africa	23.7 (24.0)	39.9 (42.1)	36.5 (34.0)	59.9 (61.7)	15.25 (15.20)
Spain	41.8 (41.8)	30.6 (30.3)	27.6 (27.9)	48.9 (48.3)	15.94 (15.97)
Turkey	36.0 (36.4)	43.8 (43.1)	19.6 (19.4)	55.7 (56.1)	16.35 (16.51)

class sizes of over 40 students were of concern, because they might have been two classes taught by the same teacher. The random selection attempted to control for this. The selection was completely random except for two criteria. A number of 355 students was not considered for selection because they had more than five missing values on the teaching behaviour or more than two missing values on engagement questionnaire. Another number of 653 students that could not be classified to one of the domains language, natural sciences or social sciences subjects.

6.2 Measures

Student perceptions of teaching behaviour was measured using My Teacher Questionnaire (Maulana et al., 2015a, b; Van der Lans et al., 2015) The questionnaire comprises 41 items that operationalize six domains of teaching behaviour: safe learning climate, clear and structured instruction, activating teaching, teaching learning strategies, differentiation. Response categories were provided on a 4-point Likert scale, ranging from 1 (never) to 4 (often). Students' responses were calibrated into a unidimensional and comparable metric using Partial Credit Model (PCM) with quasi-international concurrent linking method (Van der Lans et al., 2021). Hence, teaching behaviour was analyzed as the unidimensional construct. This unidimensional construct has been proven to be valid and reliable, as well as invariant, across the six countries (Van der Lans et al., 2021).

Perceived engagement was measured using the student engagement scale of Skinner et al. (2009). This scale measures emotional (5 items; e.g. "In this class I feel good"), and behavioural engagement (5 items; e.g. "In this class I pay attention"). Reliability of the measure is satisfactory (Cronbach's alpha = 0.81 and 0.83 for behavioural and emotional engagement respectively). All items are scored on a four-point scale, with higher responses indicating higher engagement levels. To examine the measurement invariance across the six countries, the engagement scale was subject to Multilevel Multi-Group Confirmatory Factor Analysis (MMGCFA) in which students were clustered within teachers, after the factor structure was

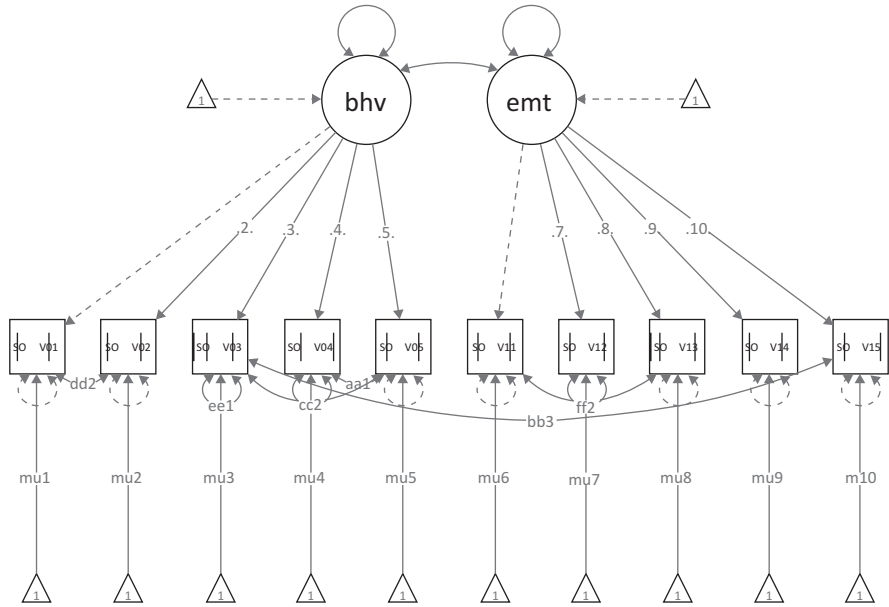


Fig. 20.1 The tested factor structure of behavioural and emotional engagement

confirmed in each country data (see Fig. 20.1). For this analysis, a random subset of five students was selected from every teacher. Results show that the scale reached the partial metric invariance⁵ (CFI = 0.99, TLI = 0.98, RMSEA = 0.08) (Hoyle & Panter, 1995), allowing us to compare the correlation between countries.

Both questionnaires were translated and back-translated for use in the six countries using procedures in accordance with International Test Commission (2017).

⁵The parameters freed were: South Africa = residual correlation item 4 and 5 was, The Netherlands = residual correlation item, Indonesia, South Korea and Spain = residual correlation item 3 and 10, item 6 and 8, Indonesia = residual correlation item 3 and 5, South Korea, the Netherlands and South Africa = the scaling factor of item 3. Overall, item 3 was the most problematic. Interpretation of partial invariance suggests that not all parameters have identical interpretation across countries. In this study, only a small number of items does not meet strict invariant. By freeing the residual correlations (and one scaling factor), we were able to fix all 10 factor loadings to be identical across countries. Since factor loadings determine the factor’s metric, comparing the factor scores and associations between countries with relatively small number of freed residual inter-item correlations is deemed acceptable. Item deletion was not preferred because all items were assumed to measure aspects of perceived teaching behaviour uniquely based on face validity, and minor violations in the model are allowed because the model is not perfect but is still within the acceptable boundary given the complexity of the model and the large number of variables and parameters involved.

6.3 Analysis Approach

To answer the first research question, “*How does the relationship between perceived teaching behaviour and student engagement compare between countries?*”, as a first step, simple Pearson product correlations were estimated correlating the mean score with behavioral engagement and emotional engagement with the estimated perceived teaching behaviour. The concurrent quasi-international calibration approach was employed (see Van der Lans et al., 2021). Then, the correlations between the latent (estimated) student engagement and perceived teaching behaviour were computed, by applying a multi-level multi-group SEM model. Student perceived teaching behaviour was added as a predictor of the latent variables student behavioral and emotional engagement.

To answer the second research question, “*How does student gender, student age, and school subject mediate the relationship between student engagement and perceived teaching behaviour across countries?*” three multi-level multi-group SEM analyses were performed. Three models were examined testing the mediation effect of student gender, student age and school subject separately. Results focus on the variation in the predictive effect of perceived teaching behaviour on students’ behavioral and emotional engagement.

All analyses were performed in R (4.0.3) (R Core Team, 2021), and Rstudio (1.1.456) and SPSS 27. The R packages used were “lavaan” (version info; author) and “SEMtools” (version info; Jorgensen, 2021). All SEM models considered the items scores on the engagement as ordinal. Estimation was performed with the diagonally weighted least squares (DWLS) estimator. Although teacher ID codes were available, lavaan does not allow multilevel estimation with ordered data. An empirical comparison using model fit coefficients favoured the specification of ordered item responses over the specification of interval level item responses. Models were grouped by country identification code (ID) using the group command. The three mediators were coded: Student gender (0 = male, 1 = female), Student age (1 = 11 years, until 10 = 20 years old), and teacher subject (1 = languages, 2 = natural sciences, 3 = social sciences,). The variable student age was considered of interval level. The variable teacher subject was dummy coded into three dummy variables, namely: subject1_dummy (0 = other subject, 1 = language), subject2_dummy (0 = other subject, 1 = natural sciences), and subject3_dummy (0 = other subject, 1 = social sciences).

7 Results

7.1 *Relationship Between Perceived Teaching Behaviour and Student Engagement*

Results show that, in general, there are differences in perceived teaching behaviour (TB) across the countries and the differences in the raw mean of student engagement scales are visible but smaller (see Table 20.2). Interestingly, South African students reported the highest raw mean of perceived behavioural engagement (BE) ($M = 3.38, SD = 0.55$) and second highest raw mean score of emotional engagement (EE) ($M = 3.30, SD = 0.57$), but the country had the second lowest (latent) mean perceived teaching behaviour (TB) score ($M = 1.90, SD = 1.50$).

In all countries, the Pearson correlations are positive, with moderate to strong in magnitude ($r = 0.32$ in Spain – 0.70 in The Netherlands). In general, the correlations are slightly stronger for emotional engagement compared to behavioural engagement. Using the latent score of engagement scales instead of the raw scores produced generally stronger correlations with latent perceived teaching behaviour in all countries (see Table 20.1). The magnitude of the correlations varies from highly moderate ($r = 0.40$ in Spain) to strong ($r = 0.80$ in The Netherlands). In general, results show that perceived teaching behaviour are related positively, and mostly strongly, to student engagement across the six countries. This means the higher the perceived teaching behaviour, the higher students reported their academic engagement, and vice versa. Slight differences in the magnitude of relationships between perceived teaching behaviour and self-reported engagement are evident. The strongest link was found in the Netherlands (0.80), followed by South Korea (0.73), South Africa, Indonesia, Turkey, and Spain.

7.2 *Student Gender, Student Age, School Subject, Teaching Behaviour and Engagement*

In the subsequent step, mediation effects were added for student gender, age and school subject separately. As a first step, the model without mediator is estimated. This model adds the regressive effect of emotional engagement and behavioural engagement separately on teaching behaviour. These direct effects are significant in all countries, except in Spain ($\beta = -0.07, p = 0.45$). In the second step, the mediation variables were added.

Regarding student gender (see Table 20.3), results reveal no mediation effects. In all countries, student gender has non-significant effect on perceived teaching behaviour. Only in Turkey and the Netherlands emotional and behavioural engagement have a significant effect on student gender, meaning that the level of engagement is related to student gender within these countries. Effects are positive for perceived

Table 20.2 Descriptives and correlations (n_{Student} = 8640)

Country	Descriptives			Pearson Product correlation		Latent score correlation (partial invariance)	
	Teaching behaviour [M (SD)]	Behavioural engagement [M (SD)]	Emotional engagement [M (SD)]	Behavioural engagement x Teaching behaviour	Emotional engagement x Teaching behaviour	Behavioural engagement x Teaching behaviour	Emotional engagement x Teaching behaviour
Indonesia	1.52 (1.24)	3.03 (0.43)	2.99 (0.46)	0.38*	0.44*	0.53*	0.54*
South Korea	3.29 (2.42)	3.32 (0.56)	3.33 (0.59)	0.64*	0.65*	0.71*	0.73*
Netherlands	1.93 (1.51)	3.04 (0.57)	2.94 (0.68)	0.58*	0.70*	0.70*	0.80*
South Africa	1.90 (1.50)	3.38 (0.55)	3.30 (0.57)	0.48*	0.48*	0.60*	0.60*
Spain	1.99 (1.10)	3.06 (0.50)	3.07 (0.57)	0.32*	0.41*	0.40*	0.49*
Turkey	2.19 (1.67)	3.04 (0.67)	2.91 (0.74)	0.42*	0.45*	0.47*	0.53*

Note. *p < 0.05

Table 20.3 Mediator Gender: Standardized direct effects of engagement on perceived teaching behaviour and its mediating indirect effect via student gender (0 = boy, 1 = girl) ($n_{\text{Student}} = 8640$)

	Direct effects		Indirect effects	
	TB × BE	TB × EE	TB × G × BE	TB × G × EE
Indonesia	0.13*	0.44*	0.004	-0.006
South Korea	0.27*	0.48*	-0.000	0.001
The Netherlands	0.21*	0.65*	0.002	-0.003
South Africa	0.28*	0.37*	0.000	-0.002
Spain	-0.08	0.55*	0.002	-0.001
Turkey	0.18*	0.38*	0.004	-0.004

Note. * $p < 0.05$; 0 = boys, 1 = girls

TB × BE = Teaching behaviour × Behavioural engagement

TB × EE = Teaching behaviour × Emotional engagement

TB × G × BE = Teaching behaviour × Gender × Behavioral engagement

TB × G × EE = Teaching behaviour × Gender × Emotional engagement

Table 20.4 Mediator Age: Standardized direct effects of engagement on perceived teaching behaviour and its mediating indirect effect via student age (continuous coding)

	Direct effects		Indirect effects	
	TB × BE	TB × EE	TB × StA × BE	TB × StA × EE
Indonesia	0.13*	0.44*	0.002	-0.002
South Korea	0.27*	0.48*	-0.002	-0.000
The Netherlands	0.21*	0.65*	-0.000	0.000
South Africa	0.28*	0.38*	0.010	-0.007
Spain	-0.09	0.56*	0.008	-0.005
Turkey	0.20*	0.36*	-0.008	0.016*

Note. * $p < 0.05$

TB × BE = Teaching behaviour × Behavioural engagement

TB × EE = Teaching behaviour × Emotional engagement

TB × StA × BE = Teaching behaviour × Student Age × Behavioral engagement

TB × StA × EE = Teaching behaviour × Student Age × Emotional engagement

behavioural ($\beta_{\text{NLD}} = 0.10$; $\beta_{\text{TR}} = 0.25$) and negative for perceived emotional engagement ($\beta_{\text{NLD}} = -0.17$; $\beta_{\text{TR}} = -0.24$).

Regarding student age (see Table 20.4), results indicate partial mediation for emotional engagement in Turkey. The direct effect of perceived emotional engagement on perceived teaching behaviour remains dominant ($\beta_{\text{TR}} = 0.36$), but in addition a small significant and positive indirect effect is found ($\beta_{\text{TR}} = 0.02$). The indirect effect suggests that relatively higher levels of perceived emotional engagement are found with relatively older students and this in turn affects the level of perceived teaching behaviour. This mediation is unique to Turkey, however, and not replicated in the other countries.

In South Korea, South Africa, Spain and Turkey, a significant direct effect of student age on perceived teaching behaviour is found. The direction, however, differs between the four countries. In Spain ($\beta_{\text{ESP}} = -0.07$) and South Africa ($\beta_{\text{ZAF}} = -0.06$), a small negative direct effect is found which indicates that older

students perceived somewhat lower levels of teaching behaviour. In Turkey ($\beta_{TR} = 0.11$) and South Korea ($\beta_{KOR} = 0.05$), a small positive direct effect is found which indicates that older students perceive somewhat higher levels of effective teaching behaviour. In the Netherlands and Indonesia, the effect of student age on teaching behaviour is non-significant. In the Netherlands, student behavioural ($\beta_{NLD} = -0.24$) and emotional engagement ($\beta_{NLD} = -0.14$) reveal a significant and negative effect on student age suggesting that student age is associated with the level of perceived engagement. This finding is unique to the Netherlands and not replicated in the other countries.

Regarding school subject (see Tables 20.5, 20.6, and 20.7), the results provide no evidence for mediation effects. In South Korea, the Netherlands, South Africa and Turkey the language subject domain has small effect on the level of perceived teaching behaviour. Only in South Korea ($\beta_{KOR} = -0.04$), the direction of the effect is negative meaning that South Korean students experienced a somewhat lower level of effective teaching behaviour in language classes. In the Netherlands ($\beta_{NLD} = 0.06$), South Africa ($\beta_{ZAF} = 0.08$) and Turkey ($\beta_{TR} = 0.05$), the students experienced a somewhat higher level of teaching in language classes. In Indonesia and Spain, no significant effects were found related to language subjects.

With respect to the domain natural sciences (see Table 20.6), the reverse pattern is observed. In South Korea, the Netherlands, South Africa and Turkey the natural science domain has small effect on the level of teaching behaviour. Only in South Korea ($\beta_{KOR} = 0.05$) the direction is positive meaning that South Korean students perceive slightly higher levels of effective teaching behaviour in natural science classes. In the Netherlands ($\beta_{NLD} = -0.04$), South Africa ($\beta_{ZAF} = -0.09$) and Turkey ($\beta_{TR} = -0.06$), the students experienced a somewhat lower level of effective teaching behaviour in language classes. In Indonesia the level of emotional engagement ($\beta_{IDN} = 0.22$) is related to a natural science subject domain, while in Spain the level of emotional ($\beta_{ESP} = 0.29$) and behavioural engagement ($\beta_{ESP} = -0.31$) significantly predicts the natural science subject domain. Within these two countries, higher levels of emotional engagement are associated with natural science classes and Spanish students report lower levels of behavioural engagement in natural science classes.

Table 20.5 Mediator Subject Language: Standardized direct effects of engagement on perceived teaching behaviour and its mediating indirect effect via school subject language (0 = No Language subject, 1 = Language subject)

	Direct effects		Indirect effects	
	TB × BE	TB × EE	TB × SubL × BE	TB × SubL × EE
Indonesia	0.13*	0.44*	-0.002	0.001
South Korea	0.27*	0.48*	-0.001	-0.001
The Netherlands	0.21*	0.65*	0.003	-0.004
South Africa	0.29*	0.37*	-0.004	0.005
Spain	-0.08	0.56*	0.000	-0.000
Turkey	0.19*	0.38*	-0.002	0.004

Note. * $p < 0.05$

Table 20.6 Mediator Subject Natural Sciences: Standardized direct effects of engagement on perceived teaching behaviour and its mediating indirect effect via school subject (0 = No Natural Sciences subject, 1 = Natural Sciences subject)

	Direct effects		Indirect effects	
	TB × BE	TB × EE	TB × Sub NS × BE	TB × Sub NS × EE
Indonesia	0.13*	0.44*	-0.007	0.005
South Korea	0.28*	0.48*	-0.005	0.004
The Netherlands	0.21*	0.65*	0.001	0.003
South Africa	0.28*	0.37*	0.004	0.001
Spain	-0.08	0.56*	0.001	-0.001
Turkey	0.19*	0.38*	-0.001	0.002

Note. * $p < 0.05$

Table 20.7 Mediator Subject Social Sciences: Standardized direct effects of engagement on perceived teaching behaviour and its mediating indirect effect via school subject (0 = No Natural Sciences subject, 1 = Natural Sciences subject)

	Direct effects		Indirect effects	
	TB × BE	TB × EE	TB × Sub SS × BE	TB × Sub SS × EE
Indonesia	0.13*	0.44*	-0.002	0.001
South Korea	0.27*	0.48*	-0.001	0.001
The Netherlands	0.21*	0.65*	-0.000	-0.004
South Africa	0.28*	0.37*	0.002	-0.001
Spain	-0.08	0.56*	0.000	-0.000
Turkey	0.19*	0.38*	0.001	-0.001

Note. * $p < 0.05$

With regard to the social science subject domain (see Table 20.7), only one significant effect is found. This effect is present in the Netherlands, indicating that higher levels of emotional engagement are evident in social science subject classes ($\beta_{NLD} = 0.14$).

8 Conclusions and Discussion

The present study aims to explore the links between perceived teaching behaviour and student engagement across six diverse national contexts. It also aims to explore the role of student gender, age, and school subject in the interplay between perceived teaching behaviour and engagement cross-nationally.

8.1 Perceived Teaching Behaviour and Student Engagement Across Countries

We found that perceived teaching behaviour and self-reported student engagement are significantly and positively related, and this finding is consistent across the six countries. Our finding suggests that perceived teaching behaviour is important for student engagement cross-nationally, and vice versa. This finding is in line with other studies showing that student perceptions of teaching behaviour have a powerful effect on students' self-report of cognitive and behavioural engagement (Bertills et al., 2019; Davidson et al., 2010).

In general, the link between perceived teaching behaviour and emotional engagement is stronger compared to behavioural engagement, and this trend is consistent across the six countries. This finding is consistent with past research in Spain (Inda-Caro et al., 2019). Self-Determination (SDT) stresses the importance of students' emotional engagement in facilitating internalization of goals, values, and important skills in schools (Ryan & Deci, 2009; Skinner & Pitzer, 2012). Students' perceived relatedness with teachers facilitates the internationalization of values and goals promoted by schools, which leads to the adoption of practices related to behavioural engagement (Ryan & Deci, 2009).

However, differences in the magnitude of correlations are visible, suggesting differential importance of perceived teaching behaviour for student engagement cross-nationally. Based on the latent score correlation, the relationship between perceived teaching behaviour and behavioural engagement is strong in Indonesia, South Korea, The Netherlands, and South Africa, and highly moderate in Spain and Turkey. For emotional engagement, strong relationships with perceived teaching behaviour are evident in all countries. Note, however, that the correlation coefficient for Spain is only very close to strong ($r = 0.49$). The relationships between perceived teaching behaviour and both types of engagement are strongest in The Netherlands, and weakest in Spain.

The Netherlands, South Korea, Spain, and Turkey can be categorized as developed contexts, while Indonesia and South Africa as developing contexts (UNDP, 2022). The fact that the weakest correlations between perceived teaching behaviour and self-reported engagement was found in Turkey and Spain suggests that the socio-economic dimension divide (e.g., developed vs. developing) plays a marginal role in explaining the link between the two psychological constructs. These differences may be more related to cultural factors of the six countries (Hofstede et al., 2010). Quantitatively (not in terms of Cohen's criteria), the link between the psychological constructs is strongest in the Netherlands, especially when it comes to perceived teaching behaviour and self-reported emotional engagement. The Netherlands has the lowest power distance index and the highest individualistic index, compared to the remaining five countries (Hofstede et al., 2010). In general, there seems to be modest between-country variation in the strength of the association. The implications of this variation need yet to be explored.

To conclude, regardless of the cultural background and the degree of country development, perceived teaching behaviour is significantly, and generally strongly, related to student engagement. The findings underscore the assumed universal importance of perceived teaching behaviour for student engagement, regardless of the cultural dimensions and the availability of the country's physical resources.

8.2 Student Gender, Age, School Subject, Teaching Behaviour and Student Engagement Across Countries

In general, the results show mixed evidence and generally do not support the presence of mediation effects. Only one mediation effect was found. Moreover, other single direct effects between perceived teaching behaviour and the mediators and between the mediators and perceived engagement were not fully consistent across countries in terms of directions and sizes. Adding the mediation variables did not substantially change the estimated direct effects of student perceived teaching behaviour on engagement. The evidence supports a view that student perceived engagement is primarily affected by student perceived teachers' teaching effectiveness and is to only a modest extent explained by student gender, age and/or school subject.

Of the tested mediator variables, student age is the only mediator variable that shows a significant association in Turkey. Mediators only have marginal effects on the link between perceived engagement and perceived teaching behaviour.

8.3 Implications

The present study contributes to the advancement of scientific knowledge by revealing the universality and some specificity of the interplay between perceived teaching behaviour, perceived engagement, and some background variables across several diverse countries. The results should be able to fill the lacuna of the scientific research in the field, at least to some extent, in WEIRD, non-WEIRD, and in between WEIRD and non-WEIRD settings with diverse cultural backgrounds, and reveal how factors in the microsystem (i.e., teaching behaviour, personal and contextual characteristics) interact in the development of student engagement in school.

The study also contributes to the measurement field to some extent. Using the latent score of engagement scales, instead of the raw scores, produced stronger correlations with latent perceived teaching behaviour in all countries. Observed variables are contaminated with measurement errors, while latent variables are stripped from measurement errors (Cole & Preacher, 2014). Ignoring measurement errors will result in an estimate of a correlation/regression coefficient that is lower than the true value (Fleiss & Shrout, 1977). Hence, correlation coefficients using latent

variables are more accurate. This finding is proven to be consistent across the six countries.

The study also has implications for educational practices. The findings of the universal importance of perceived teaching behaviour for perceived student engagement are encouraging to teachers and educators. Despite the cultural background and socio-economic development, teachers' behaviour in the classroom and how it is perceived by their students is universally important for their academic participation in school. Efforts to support teachers to improve the effectiveness of their teaching behaviour seem globally relevant given the six diverse settings in our sample. Despite the assumed universal importance, perceived teaching behaviour is linked to engagement more strongly in some countries than in other countries. This implies that the powerful impact of perceived teaching behaviour varies to some extent across different national contexts. Some countries like The Netherlands and South Korea should keep investing in the teaching quality improvement, while in other countries like Indonesia, Spain, and Turkey, investing in teaching quality improvement may need to be done in concert with other meso- and macro-level educational factors to bring students' engagement level to a higher level.

8.4 Limitations and Future Directions

Despite its strengths, the current study is subject to several limitations. First, convenience sampling was applied so generalizations of findings to country level should be handled with care until future replication studies with more representative samples are available. Furthermore, the sample and student characteristics across the five countries are not entirely similar. For example, a high proportion of inexperienced teachers dominated Dutch samples, so that perceived teaching behaviour largely applies to this teacher group. Students may perceive their teachers differently as a function of teaching experience. In the other four countries, on the contrary, higher proportions of experienced teachers were more visible. These sample characteristics may influence the results and, thus, the current results should be interpreted with caution until further replication studies with more representative samples are available.

Using self-reported engagement measure instead of actual engagement measure can also be viewed as a limitation. However, the focus of this study is on student perceptions, so the self-report measure (e.g., questionnaire) is deemed appropriate. Future research can benefit more from linking perceived teaching behaviour to actual engagement. The current study revealed the assumed universality as well as some specificity of the link between perceived teaching behaviour and perceived engagement. The findings showed that the associations between perceived teaching behaviour are significant, albeit different in magnitude. As also echoed by Lam et al. (2016), these findings underline the significance of integrating the etic and emic approaches in future cross-country studies and promote the search for the

‘middle ground’ acknowledging both cross-cultural similarities and differences (King & McInerney, 2014).

Acknowledgement We are indebted to all of our partners who have greatly contributed to this work. We would like to thank all teachers, schools, and observers participating in this large-scale study in the six countries. This work was supported by the Dutch scientific funding agency (NRO) under Grant number 405-15-732; the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea under Grant number NRF-2017S1A5A2A03067650, and the Directorate General of Higher Education of Indonesia under Grant number 04/SP2H/DRPM/LPPM-UNJ/III/2019.

References⁶

- Alrashidi, O., Phan, H. P., & Ngu, B. H. (2016). Academic engagement: An overview of its definitions, dimensions, and Major conceptualisations. *International Education Studies*, 9(12), 41–52.
- Amir, R., Saleha, A., Jelas, Z. M., Amad, R., & Hutkemri. (2014). Students’ engagement by age and gender: A cross-sectional study in Malaysia. *Middle-East Journal of Scientific Research*, 21(10), 1886–1892. <https://doi.org/10.5829/idosi.mejsr.2014.21.10.85168>
- Anderson, A. R., Christenson, S. L., Sinclair, M. F., & Lehr, C. A. (2004). Check and connect: The importance of relationships for promoting engagement with school. *Journal of School Psychology*, 42, 95–113. <https://doi.org/10.1016/j.jsp.2004.01.002>
- André, S., Maulana, R., Helms-Lorenz, M., Telli, S., Chun, S., Fernández-García, C. M., et al. (2020). Student perceptions in measuring teaching behavior across six countries: A multi-group confirmatory factor analysis approach to measurement invariance. *Frontiers in Psychology*, 11, 273.
- Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of School Psychology*, 44(5), 427–445.
- Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools*, 45(5), 369–386.
- Bell, C. A., Dobbelaer, M. J., Klette, K., & Visscher, A. (2019). Qualities of classroom observation systems. *School Effectiveness and School Improvement*, 30(1), 3–29. <https://doi.org/10.1080/09243453.2018.1539014>
- Berliner, D. C. (2004). Describing the behavior and documenting the accomplishments of expert teachers. *Bulletin of Science, Technology & Society*, 24(3), 200–212. <https://doi.org/10.1177/0270467604265535>
- Bertills, K., Granlund, M., & Augustine, L. (2019). Inclusive teaching skills and student engagement in physical education. *Front. Educ.*, 4, 74. <https://doi.org/10.3389/educ.2019.00074>
- Chiu, M. M. (2007). Families, economies, cultures, and science achievement in 41 countries: Country-, school-, and student-level analyses. *Journal of Family Psychology*, 21, 510–519. <https://doi.org/10.1037/0893-3200.21.3.510>
- Chiu, M. M., & Xihua, Z. (2008). Family and motivation effects on mathematics achievement: Analyses of students in 41 countries. *Learning and Instruction*, 18, 321–336. <https://doi.org/10.1016/j.learninstruc.2007.06.003>

⁶An earlier version of this chapter was presented at the 33rd International Congress for School Effectiveness and Improvement, January 2020, Marrakesh, Morocco.

- Chouinard, R., & Roy, N. (2008). Changes in high-school students' competence beliefs, utility value and achievement goals in mathematics. *British Journal of Educational Psychology*, 78(1), 31–50. <https://doi.org/10.1348/000709907X197993>
- Christenson, S. L., Reschly, A. L., Appleton, J. J., Berman, S., Spangers, D., & Varro, P. (2008). Best practices in fostering student engagement. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 1099–1120). National Association of School Psychologists.
- Coe, R., Aloisi, C., Higgins, S., & Major, L. E. (2014). *What makes great teaching? Review of the underpinning research* (Project Report). Sutton Trust.
- Cohen, J., Ruzek, E., & Sandilos, L. (2018). Does teaching quality cross subjects? Exploring consistency in elementary teacher practice across subjects. *AERA Open*, 4(3), 2332858418794492.
- Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods*, 19(2), 300–315. <https://doi.org/10.1037/a0033805>
- Cooper, K. S. (2014). Eliciting engagement in the high school classroom A mixed-methods examination of teaching practices. *American Educational Research Journal*, 51(2), 363–402. <https://doi.org/10.3102/0002831213507973>
- Davidson, A. J., Gest, S. D., & Welsh, J. A. (2010). Relatedness with teachers and peers during early adolescence: An integrated variable-oriented and person-oriented approach. *Journal of School Psychology*, 48(6), 483–510.
- Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01
- Den Brok, P., Brekelmans, M., & Wubbels, T. (2004). Interpersonal teacher behavior and student outcomes. *School Effectiveness and School Improvement*, 15, 407–442. <https://doi.org/10.1080/0924345051233138262>
- Den Brok, P. J., Wubbels, T., Veldman, I. M. J., & van Tartwijk, J. W. F. (2005). Teachers' self-reported and observed interpersonal behaviour in multicultural classrooms. In *Integrating multiple perspectives on effective learning environments* (pp. 1176–1177). University of Cyprus.
- Den Brok, P. D., Taconis, R., & Fisher, D. (2010). How well do science teacher do. Differences in teacher-student interpersonal behavior between science teachers and teachers of other (school) subjects. *The Open Education Journal*, 3, 44–53.
- Dotterer, A., McHale, S., & Crouter, A. (2009). The development and correlates of academic interests from childhood through adolescence. *Journal of Educational Psychology*, 101(2), 509–519. <https://doi.org/10.1037/a0013987>
- Fernández-García, C. M., Maulana, R., Inda-Caro, M., Helms-Lorenz, M., & García-Pérez, O. (2019). Student perceptions of secondary education teaching effectiveness: General profile, the role of personal factors, and educational level. *Frontiers in Psychology*, 10, 533. <https://doi.org/10.3389/fpsyg.2019.00533>
- Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research*, 59, 117–142.
- Finn, J. D. (2006). *The adult lives of at-risk students: The roles of attainment and engagement in high school* (NCES 2006–328). U.S. Department of Education, National Center for Education Statistics.
- Finn, J. D., & Cox, D. (1992). Participation and withdrawal among fourth-grade pupils. *American Educational Research Journal*, 29, 141–162.
- Fleiss, L., & Shrout, P. E. (1977). The effects of measurement errors on some multivariate procedures. *American Journal of Public Health*, 67(12), 1188–1191. <https://doi.org/10.2105/AJPH.67.12.1188>
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59–109.
- Fuller, F. F. (1969). Concerns of teachers: A developmental conceptualization. *American Educational Research Journal*, 6(2), 207–226.
- Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, 95(1), 148.

- Graeber, A., Newton, K., & Chambliss, M. (2012). Crossing the Borders again: Challenges in comparing quality instruction in mathematics and Reading. *Teachers College Record*, 114, 30.
- Grossman, P., & Stodolsky, S. (1994). Content as context: The role of school subjects in secondary school teaching. *Educational Researcher*, 24(8), 5–23.
- Guo, Y., Connor, C. M., Tompkins, V., & Morrison, F. J. (2011). Classroom quality and student engagement: Contributions to third-grade reading skills. *Frontiers in Psychology*, 2, Article 157. <https://doi.org/10.3389/fpsyg.2011.00157>
- Hadzigeorgiou, Y., & Schulz, R. M. (2019). Engaging students in science: The potential role of “narrative thinking” and “romantic understanding”. *Frontiers in Education*, 4, 38. <https://doi.org/10.3389/feduc.2019.00038>
- Hartono, F. P., Sumarno, N. U., & Puji, R. P. N. (2020). The level of student engagement based on gender and grade on history subject of senior high school students in Jember regency. *International Journal of Scientific and Technology Research*, 8, 21–26.
- Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Havik, T., & Westergard, E. (2020). Do teachers matter? Students’ perceptions of classroom interactions and student engagement. *Scandinavian Journal of Educational Research*, 64, 488–507. <https://doi.org/10.1080/00313831.2019.1577754>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, 466, 29. <https://doi.org/10.1038/466029a>
- Hidayati, S. N., & Rodliyah, R. J. (2020). Eksplorasi Strategi Guru untuk Meningkatkan Keterlibatan Siswa dalam Aktifitas Membaca [Exploring Teacher’s strategies to encourage student engagement in reading activity]. *Journal Penelitian Pendidikan*, 20(1), 121–128.
- Hofstede, G. (2001). *Culture’s consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage.
- Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: Software of the Mind* (Rev. 3 rd ed.). McGraw-Hill.
- Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 158–176). Sage.
- Inchley, J., Currie, D., Budisavljevic, S., Torsheim, T., Jastad, A., Cosma, A., et al. (2020). *Spotlight on adolescence health and wellbeing. Findings from the 2017/2018 health behaviour in school-aged children (HBSC) survey in Europe and Canada* (International Report. Volume 1. Key findings). World Health Organization.
- Inda-Caro, M., Maulana, R., Fernández-García, C. M., Peña-Calvo, J. V., Rodríguez-Menéndez, M., & Helms-Lorenz, M. (2019). Validating a model of effective teaching behaviour and student engagement: Perspectives from Spanish students. *Learning Environments Research*, 22(2), 229–251.
- International Test Commission. (2017). *The ITC guidelines for translating and adapting tests* (2nd ed.). www.InTestCom.org
- Jelas, Z. M., Salleh, A., Mahmud, I., Azman, N., Hamzah, H., Hamid, Z. A., Jani, R., et al. (2014). Gender disparity in school participation and achievement: The case in Malaysia. *Procedia – Social and Behavioural Sciences*, 140, 62–68. <https://doi.org/10.1016/j.sbspro.2014.04.387>
- Jimerson, S. R., Campos, E., & Greif, J. L. (2003). Toward an understanding of definitions and measures of school engagement and related terms. *The California School Psychologist*, 8(1), 7–27.
- Jorgensen, T. D. (2021). Package ‘semTools’. <https://cran.rproject.org/web/packages/semTools/semTools.pdf>
- King, R. B., & McInerney, D. M. (2014). Culture’s consequences on student motivation: Capturing cross-cultural universality and variability through personal investment theory. *Educational Psychologist*, 49, 175–198. <https://doi.org/10.1080/00461520.2014.926813>
- Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health*, 74(7), 262–273.

- Knapp, M. S., Shields, P. M., & Turnbull, B. J. (1995). Academic challenge in high-poverty classrooms. *Phi Delta Kappan*, 76(10), 770–776.
- Lam, S.-F., Pak, T. S., & Ma, W. Y. K. (2007). Motivating instructional contexts inventory. In P. R. Zelic (Ed.), *Issues in the psychology of motivation* (pp. 119–136). Nova Science.
- Lam, S.-F., Jimerson, S., Kikas, E., Cefai, C., Veiga, F. H., Nelson, B., et al. (2012). Do girls and boys perceive themselves as equally engaged in school? The results of an international study from 12 countries. *Journal of School Psychology*, 50(1), 77–94. <https://doi.org/10.1016/j.jsp.2011.07.004>
- Lam, S. F., Jimerson, S., Wong, B. P., Kikas, E., Shin, H., Veiga, F. H., et al. (2014). Understanding and measuring student engagement in school: The results of an international study from 12 countries. *School Psychology Quarterly*, 29(2), 213.
- Lam, S.-F., Jimerson, S., Shin, H., Cefai, C., Veiga, F. H., Hatzichristou, C., Polychroni, F., Kikas, E., Wong, B. P. H., Stanculescu, E., Basnett, J., Duck, R., Farrell, P., Liu, Y., Negovan, V., Nelson, B., Yang, H., & Zollneritsch, J. (2016). Cultural universality and specificity of student engagement in school: The results of an international study from 12 countries. *British Journal of Educational Psychology*, 86, 137–153. <https://doi.org/10.1111/bjep.12079>
- Lamote, C., Speybroeck, S., Van den Noortgate, W., & Van Damme, J. (2013). Different pathways towards dropout: The role of engagement in early school leaving. *Oxford Review of Education*, 39(6), 739–760. <https://doi.org/10.1080/03054985.2013.854202>
- Lee, V. E., & Smith, J. B. (1993). Effects of school restructuring on the achievement and engagement of middle-grade students. *Sociology of Education*, 66, 164–187.
- Lei, H., Cui, Y., & Zhou, W. (2018). Relationships between student engagement and academic achievement: A meta-analysis. *Social Behavior and Personality*, 46, 517–528. <https://doi.org/10.2224/sbp.7054>
- Levy, J., den Brok, P., Wubbels, T., & Brekelmans, M. (2003). Students' perceptions of the interpersonal aspect of the learning environment. *Learning Environments Research*, 6(1), 5–36.
- Lietaert, S., Roorda, D., Laevers, F., Verschueren, K., & De Fraine, B. (2015). The gender gap in student engagement: The role of teachers' autonomy support, structure and involvement. *British Journal of Educational Psychology*, 85, 498–518. <https://doi.org/10.1111/bjep.12095>
- Malecki, C. K., & Demaray, M. K. (2003). What type of support do they need? Investigating student adjustment as related to emotional, informational, appraisal, and instrumental support. *School Psychology Quarterly*, 18(3), 231–252. <https://doi.org/10.1521/scpq.18.3.231.22576>
- Malmberg, L., Hagger, H., Burn, K., Mutton, T., & Colls, H. (2010). Observed classroom quality during teacher education and two years of professional practice. *Journal of Educational Psychology*, 102(4), 916–932.
- Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal*, 37(1), 153–184.
- Maulana, R., & Helms-Lorenz, M. (2016). Observations and student perceptions of pre-service teachers' teaching behavior quality: Construct representation and predictive quality. *Learning Environments Research*, 19, 335–357.
- Maulana, R., Opendakker, M.-C., Den Brok, P., & Bosker, R. J. (2012). In T. Wubbels, P. van den Brok, J. van Tartwijk, & J. Levy (Eds.), *Interpersonal relationships in education: An overview of contemporary research* (pp. 207–224). Sense Publishers.
- Maulana, R., Opendakker, M., & Bosker, R. (2014). Teacher-student interpersonal relationships do change and affect academic motivation: A multilevel growth curve modelling. *British Journal of Educational Psychology*, 84(3), 459–482. <https://doi.org/10.1111/bjep.12031>
- Maulana, R., Helms-Lorenz, M., & Van de Grift, W. (2015a). Development and evaluation of a questionnaire measuring pre-service teachers' teaching behaviour: A Rasch modelling approach. *School Effectiveness and School Improvement*, 26(2), 169–194.
- Maulana, R., Helms-Lorenz, M., & Van de Grift, W. (2015b). A longitudinal study of induction on the acceleration of growth in teaching quality of beginning teachers through the eyes of

- their students. *Teaching and Teacher Education*, 51, 225–245. <https://doi.org/10.1016/j.tate.2015.07.003>
- Maulana, R., Helms-Lorenz, M., Irdiyanti, Y., & Van de Grift, W. (2016). Autonomous motivation in the Indonesian classroom: Relationship with teacher support through the lens of self-determination theory. *The Asia-Pacific Education Researcher*, 25(3), 441–451. <https://doi.org/10.1007/s40299-016-0282-5>
- Maulana, R., Helms-Lorenz, M., & Van de Grift, W. (2017). Validating a model of effective teaching behaviour of pre-service teachers. *Teachers and Teaching: Theory and Practice*, 23(4), 471–493. <https://doi.org/10.1080/13540602.2016.1211102>
- Maulana, R., André, S., Helms-Lorenz, M., Ko, J., Chun, S., Shahzad, A., Irdiyanti, Y., Lee, O., de Jager, T., Coetzee, T., & Fadhilah, N. (2021). Observed teaching behaviour in secondary education across six countries: Measurement invariance and indication of cross-national variations. *School Effectiveness and School Improvement*, 32(1), 64–95. <https://doi.org/10.1080/09243453.2020.1777170>
- Muijs, D., Kyriakides, L., Van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art—teacher effectiveness and professional learning. *School Effectiveness and School Improvement*, 25(2), 231–256.
- OECD. (2019). *PISA 2018 Results: Combined Executive Summaries Volume I, II, II*. Available at: https://www.oecd.org/pisa/Combined_Executive_Summaries_PISA_2018.pdf (09.02.2022).
- Oelsner, J., Lippold, M. A., & Greenberg, M. T. (2011). Factors influencing the development of school bonding among middle school students. *The Journal of Early Adolescence*, 31(3), 463–487. <https://doi.org/10.1177/0272431610366244>
- Oga-Baldwin, W. L. Q., & Fryer, L. (2020). Girls show better quality motivation to learn languages than boys: Latent profiles and their gender differences. *Heliyon*, 6(5). <https://doi.org/10.1016/j.heliyon.2020.e04054>
- Opendakker, M.-C., Maulana, R., & Den Brok, P. (2012). Teacher-student interpersonal relationships and academic motivation within one school year: Developmental changes and linkage. *School Effectiveness and School Improvement*, 23(1), 95–119. <https://doi.org/10.1080/09243453.2011.619198>
- Pianta, R. C., & Allen, J. P. (2008). Building capacity for positive youth development in secondary school classrooms: Changing teachers' interactions with students. In M. Shinn & H. Yoshikawa (Eds.), *Toward positive youth development: Transforming schools and community programs* (pp. 21–39). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195327892.003.0002>
- Pianta, R. C., Hamre, B. K., & Allen, J. P. (2012). Teacher-student relationships and engagement: Conceptualizing, measuring, and improving the capacity of classroom interactions. In S. L. Christenson et al. (Eds.), *Handbook of research on student engagement* (p. 365). Springer. https://doi.org/10.1007/978-1-4614-2018-7_17
- Quin, D., Hemphill, S. A., & Heerde, J. A. (2017). Associations between teaching quality and secondary students' behavioral, emotional, and cognitive engagement in school. *Social Psychology of Education*, 20(4), 807–829.
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research*, 75(2), 211–246.
- Reutzel, D. R., Child, A., Jones, C. D., & Clark, S. K. (2014). Explicit instruction in core reading programs. *The Elementary School Journal*, 114(3), 406–428.
- Richman, J. M., Rosenfeld, L. B., & Bowen, G. L. (1998). Social support for adolescents at risk of school failure. *Social Work*, 43(4), 309–323.
- Rimm-Kaufman, S. E., Baroody, A. E., Larsen, R. A. A., Curby, T. W., & Abry, T. (2015). To what extent do teacher–student interaction quality and student gender contribute to fifth graders' engagement in mathematics learning? *Journal of Educational Psychology*, 107(1), 170–185. <https://doi.org/10.1037/a0037252>

- Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher-student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research, 81*, 493–529. <https://doi.org/10.3102/0034654311421793>
- Ryan, R. M., & Deci, E. L. (2009). Promoting self-determined school engagement. In K. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 171–195). Routledge.
- Ryan, A. M., & Patrick, H. (2001). The classroom social environment and changes in adolescents' motivation and engagement during middle school. *American Educational Research Journal, 38*(2), 437–460.
- Santos, A. C., Simoes, C., Cefai, C., Freitas, E., & Arriaga, P. (2021). Emotion regulation and student engagement: Age and gender differences during adolescence. *International Journal of Educational Research, 109*. <https://doi.org/10.1016/j.ijer.2021.101830>
- Skinner, E. A., & Pitzer, J. R. (2012). Developmental dynamics of student engagement, coping, and everyday resilience. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 21–44). Springer.
- Skinner, E., Furrer, C., Marchand, G., & Kinderman, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology, 100*, 765–781. <https://doi.org/10.1037/a0028089>
- Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioural and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement, 69*(3), 493–525. <https://doi.org/10.1177/0013164408323233>
- Soenens, B., Sierens, E., Vansteenkiste, M., Dochy, F., & Goossens, L. (2012). Psychologically controlling teaching: Examining outcomes, antecedents, and mediators. *Journal of Educational Psychology, 104*(1), 108–120. <https://doi.org/10.1037/a0025742>
- Suldo, S. M., Friedrich, A. A., White, T., Farmer, J., Minch, D., & Michalowski, J. (2009). Teacher support and adolescents' subjective well-being: A mixed-methods investigation. *School Psychology Review, 38*(1), 67–85.
- Susman, E., & Rogol, A. (2013). *Puberty and psychological development*. <https://doi.org/10.1002/9780471726746.ch2>
- Telli, S. (2016). Students' perceptions of teachers' interpersonal behaviour across four different school subjects: Control is good but affiliation is better. *Teachers and Teaching, 22*(6), 729–744.
- Teoh, H. C., Abdullah, M. C., Roslan, S., & Daud, S. (2013). An investigation of student engagement in a Malaysian Public University. *Procedia – Social and Behavioural Sciences, 90*, 142–151. <https://doi.org/10.1016/j.sbspro.2013.07.075>
- Teuber, Z., Tang, X., Salmela-Aro, K., & Wild, E. (2021). Assessing engagement in Chinese upper secondary school students using the Chinese version of the schoolwork engagement inventory: Energy, dedication, and absorption (CEDA). *Frontiers in Psychology, 12*. <https://doi.org/10.3389/fpsyg.2021.638189>
- Thijs, J., & Verkuyten, M. (2009). Students' anticipated situational engagement: The roles of teacher behavior, personal engagement, and gender. *The Journal of Genetic Psychology, 170*(3), 268–286.
- Thuen, E., & Bru, E. (2000). Learning environment, meaningfulness of schoolwork and on-task-orientation among Norwegian 9th grade students. *School Psychology International, 21*(4), 393–413.
- United Nations Development Programme (UNDP). (2022). *Human development reports: Human development index (HDI)*. Retrieved from: <https://hdr.undp.org/en/indicators/137506>
- Van de Gaer, E., De Fraine, B., Pustjens, H., Van Damme, J., De Munter, A., & Onghena, P. (2009). School effects on the development of motivation toward learning tasks and the development of academic self-concept in secondary education: A multivariate latent growth curve approach. *School Effectiveness and School Improvement, 20*(2), 235–253. <https://doi.org/10.1080/09243450902883920>

- Van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research*, 49(2), 127–152.
- Van de Grift, W. J. (2014). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement*, 25(3), 295–311.
- Van de Grift, W. J. C. M., Van der Wal, M., & Torenbek, M. (2011). Ontwikkeling in de pedagogisch didactische vaardigheid van leraren in het basisonderwijs. *Pedagogische Studiën*, 88(6), 416–432.
- Van der Lans, R. M., & Maulana, R. (2018). The use of secondary school student ratings of their teacher's skillfulness for low-stake assessment and high-stake evaluation. *Studies in Educational Evaluation*, 58, 112–121.
- Van der Lans, R. M., Van de Grift, W. J., & Van Veen, K. (2015). Developing a teacher evaluation instrument to provide formative feedback using student ratings of teaching acts. *Educational Measurement: Issues and Practice*, 34(3), 18–27.
- Van der Lans, R. M., Maulana, R., Helms-Lorenz, M., Garcia, C. M. F., Chun, S., de Jager, T., Irnidayanti, J., et al. (2021). Student perceptions of teaching quality in five countries: A polytomous IRT approach to measurement invariance. *SAGE Open*. <https://doi.org/10.1177/21582440211040121>
- Vansteenkiste, M., Sierens, E., Goossens, L., Soenens, B., Dochy, F., Mouratidis, A., Aelterman, N., Haerens, L., & Beyers, W. (2012). Identifying configurations of perceived teacher autonomy support and structure: Associations with self-regulated learning, motivation and problem behavior. *Learning and Instruction*, 22(6), 431–439. <https://doi.org/10.1016/j.learninstruc.2012.04.002>
- Virtanen, T. E., Lerkkanen, M. J., Poikkeus, A. M., & Kuorelahti, M. (2015). The relationship between classroom quality and students' engagement in secondary school. *Educational Psychology*, 35(8), 963–983. <https://doi.org/10.1080/01443410.2013.822961>
- Wang, M. T., & Eccles, J. S. (2012). Social support matters: Longitudinal effects of social support on three dimensions of school engagement from middle to high school. *Child Development*, 83(3), 877–895. <https://doi.org/10.1111/j.1467-8624.2012.01745.x>
- Wang, M. T., & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement, and academic achievement in middle school. *American Educational Research Journal*, 47(3), 633–662.
- Watt, H. M. G. (2000). Measuring attitudinal change in mathematics and English over the 1st year of junior high school: A multidimensional analysis. *The Journal of Experimental Education*, 68(4), 331–361. <https://doi.org/10.1080/00220970009600642>

Ridwan Maulana is an associate professor at the Department of Teacher Education, University of Groningen, the Netherlands. His major research interests include teaching and teacher education, factors influencing effective teaching, methods associated with the measurement of teaching, longitudinal research, cross-country comparisons, effects of teaching behaviour on students' motivation and engagement, and teacher professional development. He has been involved in various teacher professional development projects including the Dutch induction programme and school–university-based partnership. He is currently a project leader of an international project on teaching quality involving countries from Europe, Asia, Africa, Australia, and America. He is a European Editor of Learning Environments Research journal, a SIG leader of Learning Environments of American Educational Research Association, and chair of the Ethics Commission of the Teacher Education.

Rikkert van der Lans obtained his PhD from the University of Groningen. He is currently a postdoctoral researcher at the University of Amsterdam. His research interest is in the application of data to inform and improve the work of professionals. Research methods are primarily quantitative. His current work spans the fields of clinical psychology and education. His current research focuses on various activities that teachers perform in schools that make the teaching profession so

beautiful and at the same time complex. He has a specific interest in making visible the wide variation in the ways teachers perform these activities. He currently works to conduct the Teaching and Learning International Survey (TALIS) 2024 for the Dutch context.

Michelle Helms-Lorenz is an Associate Professor at the Department of Teacher Education, University of Groningen, The Netherlands. Her research interest covers the cultural specificity versus universality (of behaviour and psychological processes). This interest was fed by the cultural diversity in South Africa, where she was born and raised. Michelle's second passion is education, the bumpy road toward development. Her research interests include teaching skills and well-being of beginning and pre-service teachers and effective interventions to promote their professional growth and retention. (orcid.org/0000-0001-9314-6962).

Sibel Telli is an associate professor of biology didactic at Canakkale Onsekiz Mart University, Turkey. She was the post doc researcher at University of Groningen, The Netherlands (2008-2009, funded by EU Science in Society Program) and Koblenz-Landau University, Germany (2009-2012, funded by German Research Foundation-DFG). Dr. Telli is the recipient of the Best Dissertation Award from Middle East Technical University (METU) in 2007; the Best Article in Learning Environments Research Journal SIG Learning Environments American Educational Research Association (AERA) in 2010 and Reviewer Award by Elsevier in 2017, 2021 and Wiley Journals in 2021. Dr. Telli received several fellowships and grants from national and international sources, including European Council in 2002; European Science Education Research Association (ESERA) in 2003 and Scientific and Technological Research Council of Turkey (TUBITAK) in 2006, 2007, 2013, 2016, 2020 and in 2022. She is an applied researcher and teacher educator. Her research interests is strongly interdisciplinary and focuses on learning environments research, in particular teacher-student interpersonal behaviour, teacher's behaviours and communication in the classroom; student perceptions of the learning environment; biology in science education and pre and in-service (science-biology) teacher professional development. Over the years, she has participated in a range of international and national projects and has been published in various journals. She serves on the editorial board, scientific and review committee of several national-international conferences and journals.

Yulia Irnidayanti obtained her first degree in Biology Education and PhD in Biology. She is currently a Senior Lecturer and researcher at the Biology and Biology Education Department, Universitas Negeri Jakarta [State University of Jakarta], Indonesia. Since 2001, she has been working together with the Teacher Education Department of University of Groningen, the Netherlands, on the project about teaching quality and student academic motivation from the international perspective (ICALT3/Differentiation project, Principal investigator Indonesia). She is interested in helping teachers to improve their teaching quality and student differences in their learning needs, motivation, and learning style.

Nurul Fadhilah works part-time as a lecturer at the Department of Biostatistic and Population, University of Indonesia. She has been actively involved in the international project called ICALT3/ Differentiation as an expert observer and as co-investigator for Indonesia. Currently, she is engaged in research projects related to digital health within the health informatics research cluster (HIRC). She was involved in professional teacher development for high school teachers in DKI Jakarta. She is experienced in designing and facilitating teacher professional development training, developing syllabus, designing tasks, developing differentiated instructions, especially in Cambridge IGCSE and A level Biology subject.

Carmen-María Fernández-García, PhD, Associate Professor at the Department of Educational Sciences at the University of Oviedo (Spain). She has received research grants from the Spanish Ministry of Education. She is member of the Spanish Society of Comparative Education, the Spanish Society of Pedagogy and the ASOCED Research Group. Her major research interests

involve teaching and teacher education, learning and instruction, gender and comparative education. She has published several academic papers on these topics. Currently she is joining an international project investigating teaching behavior and student outcomes across countries, the ICALT3 Project coordinated by the University of Groningen.

Mercedes Inda-Caro, PhD, Associate Professor at the University of Oviedo (Spain). She previously worked as a training support counselor in a public school as part of her FICYT scholarship training (1997) and as Child Educator for the Principality of Asturias within the Ministry of Social Services in two periods (1996/2000). Her PhD dealt with the concept of personality disorders. Currently, she is working on three lines of research: family and gender, teacher and teaching-learning education, and gender and technology studies, as a member of the ASOCED Research Group. She has several publications in scientific journals.

Prof. Seyeoung Chun is Professor Emeritus of Education at Chungnam National University, one of the major national universities in Daejeon, Korea. He received his education and Ph.D. from Seoul National University, South Korea, and has been actively engaged in education policy research and has held several key positions such as Secretary of Education to the President and CEO of KERIS. He founded the Smart Education Society in 2013, and has led many projects and initiatives for the paradigm shift of education in the digital era. Since his early career at the Korean National Commission for UNESCO, he has participated in many international cooperation projects and worked for several developing countries such as Nicaragua, Honduras, Cambodia, etc. *Education Miracle in the Republic of Korea* is the latest book to be published as a summary of his academic life.

Okhwa Lee is a professor (emeritus) at the Department of Education, Chungbuk National University, South Korea, and CEO of SmartSchool (Ltd). Okhwa Lee is a specialist in educational technology and a practitioner in pre-service teacher education. She is a pioneer of software education, e-learning, and smart education in South Korea. She has been a member of the Presidential Educational Reform Committee and the Presidential e-Government of South Korea including various department committees. She has collaborated in the global society such as European Erasmus mobility programme, UNESCO, OECD and in the Korean government ODA (Official Development Assistant) programmes for Nicaragua, Cambodia, Myanmar, Nigeria, Vietnam, Thailand, Philippines and Ethiopia.

Thelma de Jager is the HOD of the Department Educational Foundation and a senior lecturer, Tshwane University of Technology, South Africa. She received several awards for woman researcher and lecturer of the year, conducted several keynote addresses at conferences and authored and edited textbooks such as: General Subject Didactics, Creative Arts Education The Science to Teach and Differentiated Instruction. She is currently the project leader of the South African team for the ICALT 3 project and the British Council on Inclusive Education, T4ALL project. She has a passion to improve teaching pedagogy and her studies could impact education policy that speaks to the implementation of differentiated learning.

Mattheus (Thys) Coetzee is the Head of the Department Multimedia Design and Development, Tshwane University of Technology, South Africa. His research interests include teacher education, retention and success rates in engineering studies as well as statistics of student success. His fields of expertise include academic leadership in higher education, implementation of online teaching and learning technologies and the development of multimedia for teaching and learning. He was a project leader on an international project to implement eLearning in South African higher education and currently he leads an Academic Leadership project with two Dutch universities and Southern African institutions.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 21

Impact of Play-Based Pedagogies in Selected Asian Contexts: What Do We Know and How to Move Forward?



Alfredo Bautista, Jimmy Yu, Kerry Lee, and Jin Sun

Abstract In the Asian continent, many Early Childhood Education (ECE) policies have been influenced by Western theories and pedagogies. An example is the widespread presence of the notion of *play* in curriculum policy frameworks, which in part responds to research findings originated in the West. However, given what we know about cross-cultural differences in child development and learning, it is imperative to examine the state of the art on play research conducted with Asian children. This chapter reviews the literature on the impact of play-based pedagogies in Mainland China, Hong Kong, Singapore, and Japan. We describe the types of studies conducted in these jurisdictions and their overall findings, with the aim of outlining future research agendas. We describe the socio-cultural beliefs about ECE in the selected contexts and the visions of play articulated in their official policies. Then, we provide an overview of the empirical studies available, distinguishing between naturalistic and intervention studies. Studies published in English academic journals have mainly analyzed the impact of structured and guided forms of play, focusing primarily on socio-emotional outcomes, with minimal research on domains such as scientific thinking, number sense, or creativity, and no research on other areas. We argue that the existing work reflects traditional Asian values and deep-rooted beliefs about ECE, where play is seen as a rather unimportant activity. We conclude that to better justify the inclusion of play in ECE policies across Asia, it would be vital to produce an extensive, rigorous, and locally situated corpus of play impact studies.

Keywords Early childhood education · Play · Policy and practice · Child outcomes · Asia

A. Bautista (✉) · J. Yu · K. Lee · J. Sun

Department of Early Childhood Education, Centre for Educational and Developmental Sciences, The Education University of Hong Kong, Hong Kong, China

e-mail: abautista@eduhk.hk

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_21

473

1 Introduction

As a result of globalization, Early Childhood Education (ECE) policies in Asia might have been influenced by Western theories and pedagogies (Gupta, 2018; Yang & Li, 2019). One clear instance is the widespread presence of the notion of *play* in official curriculum guidelines, licensing and accreditation frameworks, and other high stakes policies, where play is commonly regarded as an essential strategy for child development and learning (Bautista, Yu, et al., 2021b; Grieshaber, 2016; Gupta, 2014). Following international trends (Organisation for Economic Co-operation and Development [OECD], 2004; The LEGO Foundation & UNICEF, 2018), Asian ECE teachers are encouraged to implement play-based pedagogies that are child-centered and process-oriented, providing children with actively engaging, meaningful, socially interactive, and joyful learning opportunities. Environments that promote play, exploration, and hands-on experiences are understood to be the core of effective ECE programs across Asia (Bautista et al., 2019; Cheung et al., 2015; Fujisawa et al., 2008; Gupta, 2014).

However, research in support of the inclusion of play in ECE settings has been primarily conducted in Western societies (Lai et al., 2018). Given what we know about significant differences in many aspects of human psychology across cultures, and particularly about the unique characteristics of the Asian learner (King & Bernardo, 2016; Li, 2010), it is imperative to examine the state of the art on play research conducted with Asian children. In view of the lack of systematic reviews in Asia, this chapter brings together the available literature on the impact of play-based pedagogies across four specific Asian contexts: Mainland China, Hong Kong, Singapore, and Japan. The chapter aims to describe the types of play impact studies conducted in these jurisdictions and their overall findings, as well as suggest future research agendas for play researchers within the Asian continent.

The chapter is structured into five sections. First, we provide a brief overview of Western research on play within ECE settings and its impact on children's development and learning. The second section describes socio-cultural beliefs about ECE in the four Asian jurisdictions considered, as well as the visions of play articulated in their official curriculum policy frameworks. In the third section, we provide an overview of the research conducted in these regions to analyze the impact of play-based pedagogies on children's developmental and/or learning outcomes, distinguishing between naturalistic and intervention studies. The fourth section critically analyzes the existing literature and identifies research gaps. Finally, the fifth section outlines future research agendas and discusses practical implications.

2 Western Research on Play and Its Impact on Children

Nowadays, official curriculum frameworks around the world (including Asian countries) suggest ECE teachers to implement different types of play-based pedagogies, in a continuum that ranges from *structured* play (activities led by teachers with

educational purposes in mind) to *free* play (activities led by children and allowing them freedom, choice, and internal agency) (Bautista et al., 2019; Hassinger-Das et al., 2017). The emphasis on play reflects the theories and pedagogies developed by influential Western authors such as Jean Piaget, Lev Vygotsky, Jerome Bruner, Carl Jung, Friedrich Fröbel and Maria Montessori, who extensively wrote about the multiple manifestations of children's play throughout the various developmental stages and/or educational levels. For example, Piaget (1962) argued that play reflects children's stages of cognitive development, starting from *functional* play (allows children to master physical actions, with or without objects), *constructive* play (children use materials to make or build something), *symbolic/fantasy* play (children invent pretend scenarios where objects or toys are used as symbols representing something else), and finally *games* (activities with pre-established rules, normally involving competition among players). Numerous taxonomies and classifications of play types have been proposed by Western scholars (for reviews, see Burghardt, 2011; Johnson et al., 2005).

Furthermore, the adoption of play-based pedagogies in Asian ECE settings might reflect the extensive body of Western research documenting the positive impacts of play on children's developmental and learning. Play impact studies in the West have utilized a variety of research methodologies (quantitative, mix-methods, qualitative) and have adopted a wide range of research designs (e.g., experimental, correlational, longitudinal, case studies). Moreover, Western scholars have documented the impact of play-based pedagogies on a variety of developmental and learning outcomes, including physical, cognitive, academic, socio-emotional, as well as mental health outcomes. For example, Western research has found a direct correlation between playful learning environments and reduced levels of obesity, heart-related problems, and chronic stress (Burdette & Whitaker, 2005). In a recent meta-analysis of 25 studies conducted in schools across Europe, Australia, United States of America (USA), Bedard et al. (2019) found that play-based and physically active classrooms may improve academic achievement and enjoyment outcomes, as compared to traditional teacher-directed schools. In a quasi-experimental study conducted in Norway with children aged 5–7, Fjortoft (2004) found that playing in a natural environment enhanced children's physical fitness, coordination, balance and agility, as children were able to play and move in landscapes that offered challenge and unpredictability.

Another large body of Western literature shows that socio-emotional competencies are best nurtured through socio-dramatic and pretend play with peers and caring adults, and other social interactions in small group settings (Yogman et al., 2018). In the USA, Pellegrini et al. (2002) conducted a longitudinal study of children's playground games, with emphasis on how play affected children's social competence and adjustment to school. It was found that facility with games predicted boys' social competence, and that play enhanced both boys' and girls' adjustment to the first year of Primary school. Finally, a large body of Western research has documented that cognitive and linguistic development are also optimized through active and exploratory forms of play (Fox et al., 2010; Liu et al., 2017a). Play enhances brain structure and promotes self-regulation and executive functioning (i.e., working memory, inhibition, shifting), which allow young children to

pursue goals and ignore distractions (Diamond, 2013). Quinn et al. (2018) conducted a meta-analysis of the literature focusing on the relationship between symbolic play and language acquisition. Drawing on 35 studies conducted in Australia, United Kingdom, Finland, and USA, the authors identified a robust association between symbolic play and language development.

3 Societal Beliefs About ECE and Curriculum Policy Visions on Play in Selected Asian Contexts

While Asia is often seen by Western scholars as a homogenous whole, the various Asian countries have specific traditions and socio-cultural characteristics (e.g., values, norms, priorities, beliefs), varied conceptions about early childhood and child development, as well as different official discourses on the role of play in ECE (Bautista, Yu, et al., 2021b; Grieshaber, 2016; Gupta, 2014). Given the lack of review studies on play conducted in Asia, this chapter focuses on four specific contexts: Mainland China, Hong Kong, Singapore, and Japan. These jurisdictions were selected due to the availability of (a) ECE policy frameworks written in English or Chinese and (b) published journal articles focusing on the impact of play-based pedagogies in young children.

Mainland China has a strong cultural tradition of placing emphasis on academic learning and achievement (Gopinathan & Lee, 2018; Li, 2010). Although this tradition is rooted in Confucianism, academic achievement in modern Chinese society is still regarded as a vehicle for social mobility. In his review on Chinese perceptions of early childhood, Luo et al. (2013) argued that this emphasis on the Confucian principle of knowledge (*Zhi*) has steered Chinese parents beliefs on learning away from avenues that entail high degrees of playfulness and enjoyment. The authors argued that other aspects of Confucian culture, in particular, the notion of *Guan* (i.e., Chinese term that means training children in the appropriate or expected behaviors) renders much of learning top-down and directed by adults. In this light, the ECE curriculum framework in Mainland China can be seen as somewhat revolutionary in its emphasis on the role of play in learning (Ministry of Education of the People's Republic of China [MOE-PRC], 2012). Indeed, the curriculum in China states that *children's learning should be derived from their "play and daily life"* and that *"we need to treasure the unique value of play"* (p.2). This strong stance on the importance of play, in its own right, is tempered in other parts of the framework where play is referred to as a tool for academic learning. For instance, the curriculum suggests that teachers let children construct their play with materials in different shapes to learn shapes and play games like drawing circles to build a foundation for writing. Suggestions such as these reveal that play is, in fact, seen as vehicle to acquire academic knowledge and skills (Li et al., 2016).

Hong Kong and Singapore are highly developed and densely populated metropolises. Both are regional trading hubs with highly developed infrastructure and world leading educational systems. Although both cities have a predominately Chinese

population and share a British colonial history, Singapore has a much larger proportion of non-Chinese in her population (~35%). Regarding parenting practices, Chinese parents in both Hong Kong and Singapore share the traditional Confucian values placed on academic achievement and tend to send their children for private tuition even before the commencement of primary school (Bull et al., 2018; Rao & Lau, 2018). In addition to cultural values, this behavior is also likely driven by parents' concerns about their children's readiness for primary education, which is often characterized as competitive and academic-oriented (Gopinathan & Lee, 2018).

Perhaps because of the immediacy of other cultures in Singapore society, the degree to which Singaporean Chinese holds on to traditional Confucian values also tends to be stronger than in Hong Kong. Although not likely an overt consideration in policy making, it is interesting to note the differences in how play is conceptualized in the curriculum frameworks of the two metropolises. Singapore's curriculum advocates purposeful play, that is, play-based pedagogies that involve activities purposefully planned by ECE teachers to achieve intended learning goals (for example, educational games, blocks, puzzles) (Singapore Ministry of Education [MOE], 2013). In contrast, Hong Kong refers to free play in her official curriculum framework, emphasizing the importance of play in drawing on or cultivating children's intrinsic motivation, autonomy, creativity, and freedom for exploration and curiosity (Curriculum Development Council [CDC], 2017).

Japan presents a very different case. In contrast to the three Chinese dominated societies examined thus far, the Japanese do not emphasize academic achievement before primary schooling. Rather, they emphasize the notion of *mimamoru* (i.e., teaching by watching and waiting), grounded in the belief of respecting children and giving children opportunities for taking up responsibility (Hayashi, 2011). Besides the hands-off approach, a key early childhood practice is group-based curriculum, which is thought to be beneficial for children's socio-emotional development (Izumi-Taylor, 2013). Rather than having direct instruction as its main function, the Japanese believe that kindergartens serve to provide opportunities for children to interact and play with others who are outside of their family circle. Echoing these societal beliefs, Japan's curriculum framework does not prescribe play for academic learning but instead focuses on the child-directed quality of play (Japan Ministry of Education, Culture, Sports, Science and Technology [MEXT], 2008). The curriculum characterizes play as a voluntary and spontaneous activity enacted by children. Play is seen as a basic form of early childhood learning. Rather than a focus on learning outcomes, the role of the teachers is to prepare an appropriate environment that corresponds to the children's play patterns and to facilitate children's engagement and enjoyment (Fujisawa et al., 2008; Takahashi, 2016).

In sum, the four selected jurisdictions have both similarities and differences in their socio-cultural beliefs about early childhood. However, play is a central component of their official ECE curriculum policy frameworks (Bautista, Yu, et al., 2021b; Grieshaber, 2016; Gupta, 2014), although with differences in the specific play approaches that teachers are encouraged to facilitate, ranging from *structured* (teacher-led) to *free* (child-led) play (Hassinger-Das et al., 2017).

4 Reviewing the Asian Literature on the Impact of Play on Child Outcomes

The key research question addressed in this section is: Drawing on the available empirical research, what do we actually know about the impact of play-based pedagogies on children's outcomes in Mainland China, Hong Kong, Singapore, and Japan? We conducted a literature search of empirical studies published up to May 2020, using the EBSCO research database. EBSCO is often used in similar review studies, given that it includes a large number of high-quality academic journals. Keywords in the search included the name of each individual Asian jurisdiction (e.g., Singapore, Japan), play, preschool OR kindergarten OR playschool, learning OR development OR outcomes, and child OR children. As this was the first systematic exploration of the topic, we decided to focus exclusively on peer-reviewed journal articles written in English.

A total of 16 articles were identified. We read the studies in detail and produced summaries highlighting the main findings. Table 21.1 presents descriptive information about the 16 articles, including publication year (ordered chronologically), jurisdiction where the study was conducted, study type, research approach employed, type of play investigated, and outcome(s) measured. Note that the category study type distinguished between *naturalistic studies* (i.e., those that explored the impact of play-based pedagogies on children within ECE programs) and *intervention studies* (i.e., those that employed controlled research designs to investigate specific outcomes of play-based pedagogies). The two following subsections further elaborate on the naturalistic and intervention studies identified, respectively.

4.1 Naturalistic Studies

Only five naturalistic studies were identified, one conducted in Hong Kong (Cheung et al., 2015), two in Singapore (Lee & Goh, 2012; Ng & Bull, 2018), and two in Japan (Fujisawa et al., 2008; Takahashi, 2016). They all showed that play-based pedagogies have the potential to positively impact specific aspects of Asian children's socio-emotional development. Note that the five studies are qualitative and based on small, non-representative samples.

English journal publications in Hong Kong and Singapore have focused on investigating the effect of guided or structured forms of play on children, at the expense of free play which has been only investigated in Japan. In Hong Kong, Cheung et al. (2015) conducted a comparative case study in two contrasting preschools. In the academically focused preschool, learning activities were organized following teachers' plans, with specified learning objectives; children were permitted to play in interest corners only after they finished the compulsory learning activities. In the play-based preschool, children were usually engaged in small-group activities and encouraged to choose their own activities, for example enjoying their

Table 21.1 Descriptive information about the 16 empirical research articles identified in the four Asian jurisdictions

Author(s)	Year	Jurisdiction	Study type	Research approach	Play type(s)	Outcome measured
Cheung	2018	Hong Kong	Intervention	Quantitative	Free and guided play	Creativity
Ng & Bull	2018	Singapore	Naturalistic	Qualitative	Outdoor play	Socio-emotional skills
Fung & Cheng	2017	Hong Kong	Intervention	Quantitative	Pretend play	Socio-emotional skills
Liu et al.	2017	Hong Kong	Intervention	Quantitative	<i>Eduplay</i>	Socio-emotional skills
Teo et al.	2017	Singapore	Intervention	Qualitative	Purposeful play	Scientific thinking
Takahashi	2016	Japan	Naturalistic	Qualitative	Pretend play	Socio-emotional skills
Li et al.	2016	Mainland China	Intervention	Quantitative	Free and guided play	Socio-emotional skills
Cheung et al.	2015	Hong Kong	Naturalistic	Qualitative	Multiple play types	Socio-emotional skills
Hui et al.	2015	Hong Kong	Intervention	Quantitative	Dramatic play	Creativity
Leung	2015	Hong Kong	Intervention	Quantitative	<i>Eduplay</i>	Socio-emotional skills
Qu et al.	2015	Singapore	Intervention	Quantitative	Sociodramatic play	Socio-emotional skills
Lee & Goh	2012	Singapore	Naturalistic	Qualitative	Pretend play	Cognitive and socio-emotional skills
Leung	2011	Hong Kong	Intervention	Quantitative	<i>Eduplay</i>	Socio-emotional skills
Wang & Hung	2010	Hong Kong	Intervention	Quantitative	Board games	Mathematics
Fujisawa et al.	2008	Japan	Naturalistic	Quantitative	Free play	Socio-emotional skills
Kok et al.	2002	Singapore	Intervention	Quantitative	Guided play	Socio-emotional skills

time in a variety of interest corners freely. Social interaction and collaborative work among children were highly encouraged. A total of 60 4–5-year-old children (30 children from each preschool) were interviewed to understand their agency orientation. Cheung et al. (2015) found that children in the academically oriented preschool had more uncertain and less participative orientation than children in the play-based preschool. The authors argued that play-based pedagogies stimulated children's capacities for agentic and participative social engagement, which in turn enhanced their chances to obtain versatile social skills. In contrast, the teacher-directed environment was more likely to undermine children's capability in expressing their own ideas and inhibiting the opportunities for children to build interactive relationships with peers and teachers.

In Singapore, Lee and Goh (2012) undertook an action research project that examined how pretend play benefited children's development and helped in their transition to primary school. Pretend play is a form of symbolic play where children use something (e.g., objects, actions ideas) to represent something else, and/or use their creativity to perform the role of imaginary characters (e.g., being superheroes, playing mummies and daddies). Children were observed as engaged in pretend play activities and comfortable to initiate activities, which echoed with newly learned knowledge. The authors concluded that young children's cognitive and affective outcomes were supported during the pretend play activities, as they were exposed to multiple opportunities to apply the knowledge learned to solve real-life problems. Through systematic natural observations in six kindergartens, Ng and Bull (2018) explored the role of teacher-child interactions in outdoor play in supporting children's social-emotional learning. The authors found that teachers provided most socio-emotional learning opportunities to children in outdoor play compared in the other three major types of learning activities (i.e., lesson time, mealtime/transition time, learning centers). During outdoor play, children were able to freely choose their activities (e.g., climbing equipment, playground play) and teachers were found to support children's interactions with peers by relating, talking and playing with peers, which facilitated relationship management and social awareness and promoted children's self-awareness and positive self-concept.

The two studies looking into the impact of free play were conducted in Japan. Fujisawa et al. (2008) investigated the reciprocity of prosocial behaviors among 3- and 4-year-old Japanese preschool children in the free play time. Two classes of 3-year-old children and two classes of 4-year-old children were observed during morning free play time for a school year. Each child was observed for 20 5-min focal observation sessions. The affiliative and prosocial behaviors occurring between the focal child and his/her peers were coded and the frequencies of each of the two types of behaviors were calculated. The results indicated positive correlations between given and received object offering and helping, as well as between the object offering and helping behaviors in the dyads. This indicated a reciprocity of prosocial behaviors during the free play time in Japanese preschool children. Findings suggest that children's prosocial behaviors can be developed and supported in positive interactions with peers during free play. In an ethnographic study, Takahashi's (2016) investigated how Japanese young children collectively constructed identities with peers in pretend play. A class with 25 children of 5 years of age in a local preschool in Japan was observed over 4 months with 8 h, 3 days a week. Based on the detailed analyses of children's conversations and interactions, three characteristic forms of interaction during play were identified, as featuring children's construction of pretend identities: (1) Reciprocal immediacy; (2) Maintaining and challenging participation; and (3) Willingness and collaboration. The author argued that play is not only for fun but implies the deliberate process of working out the roles and rules between the playmates. As a result, children co-construct their pretend identities in play situations, which contributes to support their social-emotional development (Takahashi, 2016).

4.2 Intervention Studies

Intervention studies have examined the impact on children of play-based programs and/or identified the factors influencing their effectiveness. These have been more numerous than naturalistic studies, with one study conducted in Mainland China (Li et al., 2016), seven in Hong Kong (Cheung, 2018; Fung & Cheng, 2017; Hui et al., 2015; Leung, 2011, 2015; Liu et al., 2017b; Wang & Hung, 2010), and three in Singapore (Kok et al., 2002; Qu et al., 2015; Teo et al., 2017). We did not identify research of this nature in Japan, which could be interpreted as consistent with their curriculum vision of free or unguided play (MEXT, 2008). Compared to naturalistic studies, this research has been based on more rigorous research designs, including experimental and quasi-experimental designs, with the use of both quantitative and mix-methods analytical techniques. Play interventions have been rather short in terms of duration (e.g., eight weekly sessions), typically guided or facilitated by adults (e.g., ECE teachers, parent volunteers, researchers), and implemented as extra-curricular activities.

Similar to naturalistic studies, most interventions have targeted specific outcomes related to children's socio-emotional development, with both educational and/or therapeutic purposes. For example, Liu et al. (2017a, b) showed that Hong Kong children's social competence could be improved with a parent-guided *eduplay* intervention. The notion of *eduplay* (Rao & Li, 2009) is a hybrid between the Western idea of 'playing to learn' and the Chinese Confucian emphasis on achieving outcomes pre-determined by adults. The program designed by Liu et al. (2017a, b) involved eight 1-h weekly sessions. Children engaged in collaborative group games in a classroom setting, led by trained parent volunteers. Games focused on themes related to social situations such as lining up, gathering, and dispersing. While the children were participating in the games, there were two major roles for the parent volunteers: (1) decoding social cues for children, such as summarizing the positive manners demonstrated by children during the game; (2) reinforcing children's prosocial behaviors, such as sharing and turns taking, with a star rewarding system and affirmative body language. After 8-weeks of intervention, assessed with The Early School Behavior Rating Scale (ESBRS), children's enhancement in social competence was significant based on both teacher and parent reports. The effect of the play intervention was sustainable over 5 months and generalizable to both home and ECE settings. The authors further argued that recruiting parent volunteers as instructors in play-based interventions would enhance parents' awareness and skills in facilitating children's play. In this light, parents would likely continue providing children with play opportunities and would be able to better facilitate play activities in the future.

The other classroom interventions implemented were also short and involved pretend and socio-dramatic play, which have proven effective to enhance Singaporean children's theory of mind (Qu et al., 2015) and to reduce Hong Kong children's disruptive behaviors during peer interactions (Fung & Cheng, 2017). Furthermore, interventions designed with therapeutic purposes have found that

guided forms of play contribute to reducing internalizing and externalizing behavioral problems in Hong Kong (Leung, 2011, 2015), increase time spent on social interactions in extremely shy children in Mainland China (Li et al., 2016), and enhance appropriate communication in children with autism in Singapore (Kok et al., 2002).

Beyond socio-emotional development, only three studies have analyzed the impact of play-based interventions on other child outcomes, specifically related to scientific thinking (Teo et al., 2017), creativity and problem-solving (Cheung, 2018), and mathematics (Wang & Hung, 2010). None of the intervention studies conducted in these four Asian contexts have focused on domains such as linguistic, physical, artistic, or spiritual/moral development. In the area of scientific thinking, the qualitative study by Teo et al. (2017) documented how a 90-min purposeful play session (facilitated by the researchers) allowed Singaporean children to expand their intuitive conceptions about floating and sinking. In a quasi-experimental study focusing on creativity and problem-solving, Cheung (2018) found that Hong Kong kindergarten children benefited more from a teacher-guided play approach than from a hands-off approach. In Hong Kong, Wang and Hung (2010) conducted a small-scale quasi-experimental study to examine the effect of teacher-designed boardgames on 5-year-old children's number sense. Children in the intervention group showed better number sense after 8 weekly gameplays, especially in the domain of addition-subtraction. The authors concluded that play-based pedagogies facilitate curriculum innovation and pedagogical reform, allowing ECE teachers to gain flexibility to cope with the demands of Asian parents. However, note that the small sample size was insufficient for the authors to run inferential analysis. Further studies are needed to confirm the effectiveness of this math play-based intervention.

5 Discussion

The low number of play impact studies in these four Asian contexts may be due to multiple factors. First, despite the strong advocacy of play in official curriculum policy guidelines (CDC, 2017; MEXT, 2008; MOE, 2013; MOE-PRC, 2012), play-based pedagogy is still in its infancy in many parts of Asia. In fact, except for Japan, there is a large gap between the officially sanctioned perspectives on play and the observed practices on the ground, as extensively documented in classroom-based studies (Bautista, Yu, et al., 2021b; Grieshaber, 2016; Gupta, 2014). Playtime tends to be low within many ECE settings, and it is often used instrumentally to teach about academic learning areas (Bautista et al., 2019; Lam, 2018). Contextual constraints (e.g., lack of time and space) and cultural ideologies (e.g., lack of support from school leaders, parental pressures for academic learning) are other important factors that contribute to making it difficult for ECE teachers to embrace play-based pedagogies (Bull et al., 2018; Rao & Lau, 2018). In sum, the paucity of studies may be related to the availability of ECE settings where Asian children are consistently exposed to play-based pedagogies.

This thin body of literature could be also interpreted as a manifestation of traditional Asian values and of the deep-rooted beliefs about teaching and learning, especially within Chinese societies (Gopinathan & Lee, 2018), in which play is often seen as a rather unimportant activity. Indeed, the limited work on the effects of play on domains other than socio-emotional development may be due to the Confucian belief that play is an activity with little benefit for learning, specifically for academically related learning (Luo et al., 2013). Furthermore, consistent with traditional Chinese norms, researchers have clearly favored adult-guided forms of play, also referred to as *eduplay* (Rao & Li, 2009) or purposeful play (Bautista et al., 2019), characterized by high degree of teacher structure or control, the existence of given rules, and children's lack of freedom to engage in these activities, mainly designed to achieve pre-determined outcomes. Interestingly, the essential ingredients of play (e.g., freedom, autonomy, choice, intrinsic motivation, free participation), as described by Western play theorists (e.g., Van Oers, 2013), are not visible in the studies reviewed in this chapter, except for the studies carried out in Japan.

6 Conclusion and Limitations

We conclude that Mainland China, Hong Kong, Singapore and Japan have conducted little empirical research on the impact of play-based pedagogies on children's development and learning. Taking peer-reviewed journal articles written in English as a reference, the volume of work is minimal, with only five naturalistic studies (none of them conducted in Mainland China) and 11 interventions (none of them conducted in Japan). Existing studies are small in scale, limited in scope, and often methodologically weak (i.e., short duration, focused on limited outcomes, small sample sizes, lack of control groups, lack of locally developed measures, limited generalizability).

Findings suggest that little research funding has been allocated to investigate the impact of play-based pedagogies on children in these four Asian jurisdictions. As a result of globalization in ECE (Gupta, 2018; Yang & Li, 2019), Western discourses pertaining to play seem to have been assumed as universally valid in these jurisdictions, where play-based pedagogies are recommended to teachers within ECE policies with little empirical evidence about their impact on local children (Bautista, Yu, et al., 2021b; Grieshaber, 2016; Gupta, 2014). However, we agree with J. Li (2010) in that "long-held Western assumptions about processes, efficacy, and effectiveness of learning cannot be readily applied to the study of learners from non-Western cultures" [...] because these assumptions "were developed by Western researchers to study Western people based on Western cultural norms and values" (p. 42). In other words, what is known from Western research about how play impacts on Western children may or may not be applicable to children in other parts of the world, including Asia, as cultural contexts lead to significant differences in developmental and learning pathways (UNESCO, 2010). A given play-based pedagogy that

is effective (and culturally appropriate) in the West may or may not be effective (or culturally appropriate) in the East (Bautista, Bull, et al., 2021a; Gupta, 2014).

One obvious limitation of this review is that we only included articles published in English. Nevertheless, compared to the vast volume of Western work in this area, it seems clear that there is a need for a more solid corpus of play-based research in these four Asian societies; research that takes into consideration their socio-cultural characteristics and the developmental pathways of local young children (King & Bernardo, 2016; Li, 2010). We propose future lines of research and implications in the following section.

7 Future Research and Implications

To better justify the inclusion of play and play-based pedagogies within Asian ECE curriculum frameworks, we claim it would be vital to conduct more rigorous and ambitious impact studies. Long-term longitudinal projects, which track children educated in various types of ECE settings (from academically oriented to play-based), are needed to understand the extent to which exposure to play in ECE makes a difference in the life of children (Cheung et al., 2015; Fung & Cheng, 2017). Consistent with the vision of holistic and balanced development (e.g., CDC, 2017; MOE, 2013), a wide range of outcomes should be investigated in these studies (e.g., physical, socio-emotional, cognitive, academic, mental health). Large-scale intervention studies, including randomized controlled trials, should be also undertaken to examine the benefits of specific play pedagogies within Asian ECE settings (Bull & Bautista, 2018; Cheung, 2018). In particular, it would be vital to examine the impact of different types of play in the continuum from *structured* (teacher-led) to *free* (child-led) play. As argued by Fung and Cheng (2017), studies on gender differences in response to diverse play approaches would be also desirable. Following the example of Western scholars, a wide range of research methodologies (quantitative, mix-methods, qualitative) and research designs (e.g., correlational, longitudinal, case studies) should be employed.

Developing this future research agenda would be vital not only to better justify the inclusion of play in curriculum frameworks, but also to influence (and eventually change) societal mindsets about the importance of play among Asian parents, who often prioritize academic work, discipline, and effort over other forms of learning (Lam, 2018; Rao & Li, 2009). An extensive, rigorous, and locally situated corpus of play impact studies would allow them to choose the best ECE for their children, within the frame of their respective cultural contexts (Bull et al., 2018; Rao & Lau, 2018; Yang & Li, 2019).

Acknowledgements This study was supported by the Central Reserve Fund at The Education University of Hong Kong, as part of the project “A multi-disciplinary research program in research on child development” (04A05). The views expressed in this paper are the authors’ and do not necessarily represent the views of their respective institutions.

References

- Bautista, A., Habib, M., Eng, A., & Bull, R. (2019). Purposeful play during learning center time: From curriculum to practice. *Journal of Curriculum Studies*, 51(5), 715–736. <https://doi.org/10.1080/00220272.2019.1611928>
- Bautista, A., Bull, R., Ng, E. L., & Lee, K. (2021a). “That’s just impossible in my kindergarten.” Advocating for ‘glocal’ early childhood curriculum frameworks. *Policy Futures in Education*, 19(2), 155–174. <https://doi.org/10.1177/1478210320956500>
- Bautista, A., Yu, J., Lee, K., & Sun, J. (2021b). Play in Asian preschools? Mapping a landscape of hindering factors. *Contemporary Issues in Early Childhood*, 22(4), 312–327. <https://doi.org/10.1177/14639491211058035>
- Bedard, C., St John, L., Bremer, E., Graham, J. D., & Cairney, J. (2019). A systematic review and meta-analysis on the effects of physically active classrooms on educational and enjoyment outcomes in school age children. *PLoS One*, 14(6), e0218633. <https://doi.org/10.1371/journal.pone.0218633>
- Bull, R., & Bautista, A. (2018). A careful balancing act: Evolving and harmonizing a hybrid system of ECEC in Singapore. In S. L. Kagan (Ed.), *The early advantage: Early childhood systems that lead by example* (pp. 155–181). Teachers College Press.
- Bull, R., Bautista, A., Salleh, H., & Karuppiah, N. (2018). *A case study of the Singapore early childhood education and care system. Evolving a harmonized hybrid system of ECEC: A careful balancing act*. Teachers College Press. Report commissioned by USA National Center on education and the economy (NCEE). Retrieved from: [ncee.org/wp-content/uploads/2019/03/EA-Singapore-Case-Study-022819.pdf](https://www.ncee.org/wp-content/uploads/2019/03/EA-Singapore-Case-Study-022819.pdf).
- Burdette, H., & Whitaker, R. (2005). Resurrecting free play in young children: Looking beyond fitness and fatness to attention, affiliation, and affect. *Archives of Pediatrics & Adolescent Medicine*, 159, 46–50.
- Burghardt, G. M. (2011). Defining and recognizing play. In A. D. Pellegrini (Ed.), *Oxford handbook of the development of play* (pp. 9–18). Oxford University Press.
- Cheung, R. (2018). Play-based creativity-fostering practices: The effects of different pedagogical approaches on the development of children’s creative thinking behaviours in a Chinese preschool classroom. *Pedagogy, Culture & Society*, 26(4), 511–527. <https://doi.org/10.1080/14681366.2018.1424725>
- Cheung, D., Reunamo, J., Cooper, P., Liu, K., & Vong, K.-i. P. (2015). Children’s agentive orientations in play-based and academically focused preschools in Hong Kong. *Early Child Development and Care*, 185(11–12), 1828–1844. <https://doi.org/10.1080/03004430.2015.1028400>
- Curriculum Development Council (CDC). (2017). *Kindergarten education curriculum guide: Joyful learning through play, balanced development all the way*. https://www.edb.gov.hk/attachment/en/curriculum-development/major-level-of-edu/preprimary/ENG_KGECG_2017.pdf
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168. <https://doi.org/10.1146/annurev-psycho-113011-143750>
- Fjortoft, I. (2004). Landscape and play: The effects of natural environments on children’s play and motor development. *Children, Youth and Environments*, 14(2), 21–44.
- Fox, S. E., Levitt, P., & Nelson, C. A. (2010). How the timing and quality of early experiences influence the development of brain architecture. *Child Development*, 81(1), 28–40. <https://doi.org/10.1111/j.1467-8624.2009.01380.x>
- Fujisawa, K., Kutsukake, N., & Hasegawa, T. (2008). Reciprocity of prosocial behavior in Japanese preschool children. *International Journal of Behavioral Development*, 32(2), 89–97. <https://doi.org/10.1177/0165025407084055>
- Fung, W.-K., & Cheng, R. (2017). Effect of school pretend play on preschoolers’ social competence in peer interactions: Gender as a potential moderator. *Early Childhood Education Journal*, 45(1), 35–42. <https://doi.org/10.1007/s10643-015-0760-z>

- Gopinathan, S., & Lee, M. (2018). Excellence and equity in high performing education systems: Policy lessons from Singapore and Hong Kong. *Infancia y Aprendizaje*, 41(2), 203–247. <https://doi.org/10.1080/02103702.2018.1434043>
- Grieshaber, S. (2016). Play and policy in early childhood education in the Asia-Pacific region. *Asia-Pacific Journal of Research in Early Childhood Education*, 10, 7–28. <https://doi.org/http://www.pecerajournal.com/?page=5&a=10753592>
- Gupta, A. (2014). *Diverse early childhood education policies and practices. Voices and Images from five countries in Asia*. Routledge.
- Gupta, A. (2018). How neoliberal globalization is shaping early childhood education policies in India, China, Singapore, Sri Lanka and the Maldives. *Policy Futures in Education*, 16(1), 11–28. <https://doi.org/10.1177/1478210317715796>
- Hassinger-Das, B., Toub, T. S., Zosh, J. M., Michnick, J., Golinkoff, R., & Hirsh-Pasek, K. (2017). More than just fun: A place for games in playful learning. *Journal for the Study of Education and Development*, 40(2), 191–218. <https://doi.org/10.1080/02103702.2017.1292684>
- Hayashi, A. (2011). The Japanese hands-off approach to curriculum guidelines for early childhood education as a form of cultural practice. *Asian-Pacific Journal of Research in early childhood education*, 5(2), 107–123.
- Hui, A. N., Chow, B., Chan, A., Chui, B., & Sam, C. (2015). Creativity in Hong Kong classrooms: Transition from a seriously formal pedagogy to informally playful learning. *Education 3–13*, 43(4), 393–403. <https://doi.org/10.1080/03004279.2015.1020652>
- Izumi-Taylor, S. (2013). Scaffolding in group-oriented: Japanese preschools. *YC Young Children*, 68(1), 70–75.
- Japan Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2008). Course of study for kindergarten. Ministry of Education, Culture, Sports, Science and Technology. Retrieved from https://www.mext.go.jp/component/english/_icsFiles/afieldfile/2011/04/07/1303755_002.pdf.
- Johnson, J. E., Christie, J. F., & Wardle, F. U. (2005). *Play, development, and early education*. Pearson.
- King, R., & Bernardo, A. (Eds.). (2016). *The psychology of Asian learner: A festschrift in honor of David Watkins*. Springer.
- Kok, A., Kong, T. Y., & Bernard-Opitz, V. (2002). A comparison of the effects of structured play and facilitated play approaches on preschoolers with autism: A case study. *Autism*, 6(2), 181–196. <https://doi.org/10.1177/1362361302006002005>
- Lai, N., Ang, T., Por, L., & Liew, C. (2018). The impact of play on child development: A literature review. *European Early Childhood Education Research Journal*, 26(5), 625–643. <https://doi.org/10.1080/1350293X.2018.1522479>
- Lam, P. (2018). *Bridging beliefs and practices: A study of Hong Kong kindergarten teachers' perceptions of "learning through play" and the implementation of "play" in their practices*. ProQuest Dissertations Publishing.
- Lee, S., & Goh, G. (2012). Action research to address the transition from kindergarten to primary school: Children's authentic learning, construction play, and pretend play. *Early Childhood Research & Practice*, 14(1), 1–24.
- Leung, C.-H. (2011). An experimental study of eduplay and social competence among preschool students in Hong Kong. *Early Child Development and Care*, 181(4), 535–548. <https://doi.org/10.1080/03004431003611487>
- Leung, C.-H. (2015). Enhancing social competence and the child-teacher relationship using a child-centred play training model in Hong Kong preschools. *International Journal of Early Childhood*, 47(1), 135–152. <https://doi.org/10.1007/s13158-014-0117-6>
- Li, J. (2010). Learning to self-perfect: Chinese beliefs about learning. In C. Chan & N. Rao (Eds.), *Revisiting the Chinese learner* (pp. 35–69). Springer.
- Li, Y., Coplan, R., Wang, Y., Yin, J., Zhu, J., Gao, Z., & Li, L. (2016). Preliminary evaluation of a social skills training and facilitated play early intervention programme for extremely shy young children in China. *Infant and Child Development*, 25(6), 565–574. <https://doi.org/10.1002/icd.1959>

- Liu, C., Solis, L., Jensen, H., Hopkins, E., Neale, D., Zosh, J., Hirsh-Pasek, K., & Whitebread, D. (2017a). *Neuroscience and learning through play: A review of the evidence*. The LEGO Foundation.
- Liu, S., Yuen, M., & Rao, N. (2017b). A play-based programme (pillars of society) to foster social skills of high-ability and average ability primary-one students in Hong Kong. *Gifted Education International*, 33(3), 210–231. <https://doi.org/10.1177/0261429415581221>
- Luo, R., Tamis-LeMonda, C. S., & Song, L. (2013). Chinese parents' goals and practices in early childhood. *Early Childhood Research Quarterly*, 28(4), 843–857. <https://doi.org/10.1016/j.ecresq.2013.08.001>
- Ministry of Education of the People's Republic of China (MOE-PRC). (2012). *Early learning and development guideline age 3–6*. Retrieved from <http://library.ttcw.com/uploadfiles/wyszsg/1461663255.pdf>
- Ng, S.-C., & Bull, R. (2018). Facilitating social emotional learning in kindergarten classrooms: Situational factors and teachers' strategies. *International Journal of Early Childhood*, 50(3), 335–352. <https://doi.org/10.1007/s13158-018-0225-9>
- Organisation for Economic Co-operation and Development (OECD). (2004). *Starting strong: Curricula and pedagogies in early childhood education and care: Five curriculum outlines*. OECD Publishing. Retrieved from <http://www.oecd.org/edu/school/31672150.pdf>
- Pellegrini, A., Kato, K., Blatchford, P., & Baines, E. (2002). A short-term longitudinal study of children's playground games across the first year of school: Implications for social competence and adjustment to school. *American Educational Research Journal*, 39(4), 991–1015. <https://doi.org/10.3102/00028312039004991>
- Piaget, J. (1962). *Play, dreams and imitation in childhood*. Norton.
- Qu, L., Shen, P., Chee, Y. Y., & Chen, L. (2015). Teachers' theory-of-mind coaching and children's executive function predict the training effect of sociodramatic play on children's theory of mind. *Social Development*, 24(4), 716–733. <https://doi.org/10.1111/sode.12116>
- Quinn, S., Donnelly, S., & Kidd, E. (2018). The relationship between symbolic play and language acquisition: A meta-analytic review. *Developmental Review*, 49, 121–135. <https://doi.org/10.1016/j.dr.2018.05.005>
- Rao, N., & Lau, C. (2018). Responsive policymaking and implementation: Enhancing quality and equity in Hong Kong's ECEC system. In S. L. Kagan (Ed.), *The early advantage 1—Early childhood systems that lead by example* (pp. 99–127). Teachers College Press.
- Rao, N., & Li, H. (2009). "Eduplay": Beliefs and practices related to play and learning in Chinese kindergartens. In I. Pramling-Samuelsson & M. Flear (Eds.), *Play and learning in early childhood settings* (Vol. 1, pp. 97–116). Springer. https://doi.org/10.1007/978-1-4020-8498-0_5
- Singapore Ministry of Education (MOE). (2013). *Nurturing early learners: A curriculum for kindergartens in Singapore. Educators' guide: Overview*. Ministry of Education. Retrieved from <https://www.moe.gov.sg/docs/default-source/document/education/preschool/files/nel-edu-guide-overview.pdf>
- Takahashi, M. (2016). Consumers at play: Negotiations of identity in a Japanese preschool. *Young Consumers*, 17(1), 94–106. <https://doi.org/10.1108/YC-08-2015-00549>
- Teo, T.-W., Yan, Y.-K., & Ong, M. (2017). An investigation of Singapore preschool children's emerging concepts of floating and sinking. *Pedagogies*, 12(4), 325–339. <https://doi.org/10.1080/1554480X.2017.1374186>
- The LEGO Foundation & UNICEF. (2018). *Learning through play. Strengthening learning through play in early childhood education programmes*. Education section, Programme division. United Nations Children's Fund (UNICEF).
- United Nations Educational Scientific and Cultural Organization (UNESCO). (2010). *Culture and development. Evolution and prospects*. UNESCO.
- Van Oers, B. (2013). Is it play? Towards a reconceptualisation of role play from an activity theory perspective. *European Early Childhood Education Research Journal*, 21(2), 185–198.
- Wang, Z., & Hung, L. (2010). Kindergarten children's number sense development through board games. *International Journal of Learning*, 17(8), 19–31. <https://doi.org/10.18848/1447-9494/CGP/v17i08/47181>

- Yang, W., & Li, H. (2019). *Early childhood curriculum in the era of globalisation. Policies, practices, and prospects*. Routledge.
- Yogman, M., Garner, A., Hutchinson, J., Hirsh-Pasek, K., & Golinkoff, R. M. (2018). The power of play: A pediatric role in enhancing development in young children. *Pediatrics*, *142*(3), e20182058. <https://doi.org/10.1542/peds.2018-2058>

Alfredo Bautista is Associate Professor and Associate Head of the Department of Early Childhood Education (ECE) at The Education University of Hong Kong (EdUHK). He is also Co-Director of the Centre for Educational and Developmental Sciences (CEDS). Alfredo is interested in curriculum, pedagogy, instructional practices, and teacher learning and professional development, with emphasis on the field of ECE. His work has been published in prestigious academic journals. Alfredo serves as Editor-in-Chief for the Journal for the Study of Education and Development, and Associate Editor for other academic journals. Address: Department of Early Childhood Education. B3–2/F-34 | 10 Lo Ping Road, Tai Po, New Territories. Hong Kong SAR (China). email: abautista@eduhk.hk

Jimmy Yu is a Research Assistant at CEDS at EdUHK. He works for “The PACE Project: A multi-disciplinary research program in research on child development”, which investigates preschoolers’ multi-disciplines developmental trajectories in Hong Kong. Jimmy graduated in Psychology from Shue Yan University. He is interested in investigating effective and joyful learning strategies for children through evidence-based practices. He is also interested in the neuroscience of morality, how moral sense emerges in young children, and the nature of their ethical decisions. Address: Department of Early Childhood Education. B3–2/F-34 | 10 Lo Ping Road, Tai Po, New Territories. Hong Kong SAR (China).

Kerry Lee is Professor, Head of the ECE Department, and Director of the CEDS at EdUHK. Professor Lee is trained as a cognitive developmental psychologist. His work focuses on mathematical achievement, working memory, and the development of executive functioning. He serves on the editorial or review boards of several academic journals. Address: Department of Early Childhood Education. B2–1/F-36 | 10 Lo Ping Road, Tai Po, New Territories. Hong Kong SAR (China).

Jin Sun is Assistant Professor and Associate Head of ECE Department at EdUHK, as well as Co-Director for CEDS. Her research interests include the assessment of early learning and development, early self-regulation and math development, and Chinese socialization. She is particularly interested in early development and education of socially and economically disadvantaged children. Dr. Sun has undertaken consultancies for UNICEF, UNESCO, Education for All, and the Plan. Address: Department of Early Childhood Education. B3–1/F-16 | 10 Lo Ping Road, Tai Po, New Territories. Hong Kong SAR (China).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 22

Effective Interpersonal Relationships: On the Association Between Teacher Agency and Communion with Student Outcomes



Perry den Brok, Jan van Tartwijk, and Tim Mainhard

Abstract This chapter reviews research that has investigated the link between teacher-student interpersonal relationships and student outcomes. First, prior research reviews investigating the relationship between these two sets of variables is discussed. Such research overwhelmingly shows the importance of warm and supportive relationships for both cognitive and affective outcomes, with affective outcomes also acting as an intermediary between the other two variables. Next, interpersonal theory is discussed, that conceptualizes interpersonal relationships from a systems perspective and distinguishes between the communion and agency dimensions of relationships. At the end of the contribution, research is reviewed that has used interpersonal theory as its leading framework and that has mapped students' perceptions of interpersonal relationships with one particular instrument, the Questionnaire on Teacher Interaction (QTI). Findings show that both interpersonal dimensions are positively related to cognitive as well as affective outcomes, either jointly or separately, with agency being more strongly related to cognitive outcomes and communion being more strongly related to affective outcomes.

Keywords Teacher-student interpersonal relationships · Agency · Communion · Student outcomes · Questionnaire on teacher interaction (QTI)

P. den Brok (✉)

Education and Learning Sciences, Wageningen University and Research,
Wageningen, The Netherlands

e-mail: perry.denbrok@wur.nl

J. van Tartwijk

Graduate School for Teaching, Utrecht University, Utrecht, The Netherlands

T. Mainhard

Institute of Education and Child Studies, Leiden University, Leiden, The Netherlands

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_22

1 Introduction

A vast amount of research has shown that the learning environment directly and indirectly influences students' learning and learning outcomes (Fraser, 2014). As part of the learning environment, the teacher is one of the most important factors in determining students' learning processes (Hattie, 2009). Teachers influence students in several ways, such as via providing assignments and homework, assessment, contact with parents and other teachers, and by providing instructional, emotional, and other support. Through their teaching, teachers seem to affect students' time on task (Fraser et al., 1987), emotional security (Thijs & Koomen, 2008), beliefs in their learning potential (Muijs et al., 2014), motivation and engagement (Martin & Dawson, 2009), and peer interaction (Hughes et al., 2001).

The present chapter focuses on a specific aspect of teaching in the classroom: teacher-student relationships. According to Roorda et al. (2017; also see Cornelius-White, 2007) a beneficial teacher-student relationship stimulates learning and helps to create a safe, positive classroom climate. Negative teacher-student relationships, on the other hand, may lead to feelings of insecurity and may make it harder for students to meet the demands of the school context. Also, interpersonal relationships are seen as one of the main factors in classroom management, and as such conditional to other elements in teaching and the learning environment (Evertson & Weinstein, 2006; Fraser et al., 1987; van der Lans et al., 2020).

In this contribution, we discuss teacher-student communication in terms of interpersonal theory. Interpersonal theory conceptualizes this communication in terms of two dimensions: communion or interpersonal warmth; and agency or influence (Wubbels et al., 2006). Agency refers to the degree to which someone, in this case the teacher, is perceived as dominant in or control in an interpersonal interaction; communion refers to the degree to which someone is perceived as empathic, social, harmonious or friendly (Gurtman, 2009).

The aim of the narrative review in this chapter is to investigate (1) if and to what degree both interpersonal dimensions are related to (cognitive and affective) student outcomes, and (2) to see to what degree these associations can be found in different countries and contexts across the world. In doing so, this review adds to existing reviews in several ways.

First, most of the existing research investigating links between interpersonal relationships and student outcomes focuses on just one of the two relational dimensions, such as research departing from frameworks such as self-determination (Ryan & Deci, 2000), approach-avoidance theory (Witt et al., 2004), engagement theory (Roorda et al., 2011) or student-centered relational theory (Cornelius-White, 2007), most of which focus on the communion dimension (see also Sect. 2). While there is a large number of studies in the domain of classroom management investigating the role of rules, behavior interventions by teachers or teacher punishment (e.g. Evertson & Weinstein, 2006), these studies do not relate such aspects of teaching to one (or both) of the potentially underlying interpersonal dimensions and as such research

on the influence or agency dimension in relation to student outcomes is limited (see Sect. 2).

Second, the present review uses a set of studies that all depart from the same theoretical framework (the interpersonal circumplex; Leary, 1957), focus on student perceptions of the relationship rather than a variety of methods also including observations and teacher perceptions, and use the same instrument to map these perceptions, namely the Questionnaire on Teacher Interaction (Wubbels et al., 2006; see also Sect. 3). This enhances the comparability and interpretation of the various studies discussed.

Third, as communication and perceptions are influenced by contextual and cultural factors such as values and beliefs with respect to for example individualism versus collectivism or attitudes towards leadership (den Brok & van Tartwijk, 2015), it is interesting to see whether students in different countries have different perceptions of the interpersonal relationships with their teachers and whether these perceptions affect student outcomes to the same degree.

2 General Evidence for the Association Between Teacher-Student Relationships and Student Outcomes

Jeffrey Cornelius-White (2007) conducted a meta-analysis on studies investigating the link between learner-centred teacher-student relationships and student outcomes. He defined such relationships as ‘empathic (understanding), unconditional positive regard (warm), genuine (self-awareness), non-directive (student-initiated and student-regulated) and encouraging’ (p. 113). His synthesis included 119 studies from 1948 to 2004 and covered primary, secondary and higher education. He found an overall average (corrected) correlation of .39 between such positive teacher-student relationships and student outcomes. He also found a slightly higher correlation with affective outcomes than with cognitive outcomes ($r = .35$ vs. $r = .31$). Moreover, highest correlations were found in studies using observational methods ($r = .40$), followed by studies using student perceptions ($r = .33$) and studies using a composite of different methods (.27). Studies using teacher perceptions produced the lowest correlations ($r = .17$).

Witt et al. (2004) reported a meta-analysis on studies investigating the link between teacher immediacy (the degree to which people approach each other based on similar cues of non-verbal and verbal behaviour) and student learning. They ground the ‘immediacy principle’ in the approach-avoidance theory that was developed in research on nonverbal behavior, suggesting that “people approach what they like and avoid what they don’t like” (Mehrabian, 1981, p. 22). Witt et al.’s meta-analysis included 93 studies from 1979 to 2001 investigating links between verbal and non-verbal immediacy on the one hand and cognitive (as measured via achievement tests), affective (as measured via motivation surveys) outcomes and self-perceived learning behaviour on the other. Their meta-analyses included mainly

studies conducted in higher education contexts, although a small number from other contexts was included as well. They found relatively high average correlations with affective outcomes or self-perceived learning ($r = .49$ to $r = .51$) but a markedly lower average correlation for cognitive learning outcomes ($r = .11$). Moreover, they found a higher average correlation for studies using perception scores via questionnaires ($r = .54$), than for studies using an experimental or observational design ($r = .31$).

Roorda et al. (2011) conducted a meta-analytic review to investigate the associations between positive and negative teacher–student relationships and students’ school engagement and achievement. Results were based on 99 studies from preschool to high school. Overall, medium to large associations were found for positive teacher–students relationships (e.g., closeness, involvement, relatedness, emotional support, warmth, and acceptance) with engagement, whereas small to medium associations were found with cognitive outcomes. Overall, the effects of negative relationships (e.g., conflict, rejection, role strain, verbal abuse, and relational negativity) were stronger in primary than in secondary education. In a follow-up meta-analysis, Roorda and colleagues (2017) investigated whether engagement acted as mediator in the association between teacher–student relationships and students’ cognitive outcomes. A total of 189 studies were included from preschool to high school. Meta-analytic structural equation modelling showed that both positive and negative relationships with achievement were partially mediated by student engagement.

Thus, overall, these review studies suggested that warm and caring relationships of teachers have an effect on both students’ cognitive and affective outcomes. The effects seem to be slightly stronger for affective outcomes than for cognitive outcomes – with the former acting as mediator. Interestingly, the reviews also seem to indicate that studies that have used students’ perceptions of teacher–student relationships find equally strong, if not stronger associations with student outcomes, than studies using other approaches to map teacher–student relationships. As such, the review in the present study, focusing on student perceptions of the teacher–relationship, can be considered relevant, as student perceptions are typically relatively easy to collect, reliable and valid (Fraser, 2014).

Interestingly, only a few review studies could be found reporting on concepts related to the teacher authority or interpersonal agency dimension and its potential relation to student outcomes, and evidence from these studies is less decisive than for the communion dimension.

Judith Pace and Anette Hemmings (2007) provided an overview of theoretical approaches to classroom authority – which can be seen as conceptually related to the agency dimension. They concluded that authority plays an important role in student compliance, student behaviour and student learning.

Schrodt et al. (2008) provided an overview of research investigating links between teachers’ use of power in the classroom and student outcomes. Similar to the conceptualisation in interpersonal theory, they regard power as ‘social influence’ in the classroom and distinguish it from teacher behaviour aimed at promoting interpersonal ties with students in the classroom. They reported that research

suggested that pro-social forms of power, (e.g. power based on expertise, support and rewards) rather than other types of power are positively associated with student ratings of their teachers, student behaviour and student outcomes.

Woolfolk Hoy and Weinstein (2006) reviewed research on teacher and student perceptions of teacher classroom management and concluded that a host of studies suggest that warm and demanding teachers succeed best in stimulating their classes to high achievement and cognitive outcomes. They argue that demanding or authoritative behaviour is important for student outcomes, yet in combination with warmth or cooperative behaviour.

3 Interpersonal Theory as Framework for Teacher-Student Relationships

3.1 Interpersonal Theory and Its Assumptions

In the remainder of this chapter, interpersonal theory will be the central focus to discuss associations between teacher-student relationships and student outcomes. Interpersonal theory highlights the importance of warmth and agency in teacher-student relationships and research has indicated the conditional nature of relationships on other processes in the classroom (Zijlstra et al., 2013). Many classroom studies based on Interpersonal theory, focused on teacher-student relationships as assessed by students' generalized perceptions of teachers' interpersonal classroom behaviours rather than focussing on dyadic relationships between a teacher and a single student.

A key assumption in interpersonal theory is that people mutually influence each other's behaviour and perceptions thereof (Strack & Horowitz, 2011). Student perceptions of their teachers' interpersonal style are the data source in the studies reviewed in the present chapter, which can be regarded as the generalized interpersonal meanings that students attach to their interactions with teachers, which are indicative of the perception of the relationship with their teacher (cf. Wubbels et al., 2006, 2014). These perceptions of the relationships originate in moment-to-moment verbal and nonverbal interactions (Granic & Hollenstein, 2003; van Tartwijk, 1993; Watzlawick et al., 1967); however, these moment-to-moment interactions are not the focus of the present review. Since both students and the teacher mutually influence each other, searching for causes of either healthy or problematic communication by looking at only one of these two sides is usually not productive (e.g., Watzlawick et al., 1967; Wubbels et al., 1988).

Another important assumption within this theory is that all behaviours of people, or perceptions thereof, can be described with two dimensions that together form a circumplex structure (Sadler et al., 2011): agency and communion (see Fig. 22.1). As indicated earlier, agency refers to the degree the teacher is perceived as dominant in or control; communion refers to the degree to which the teacher is perceived as

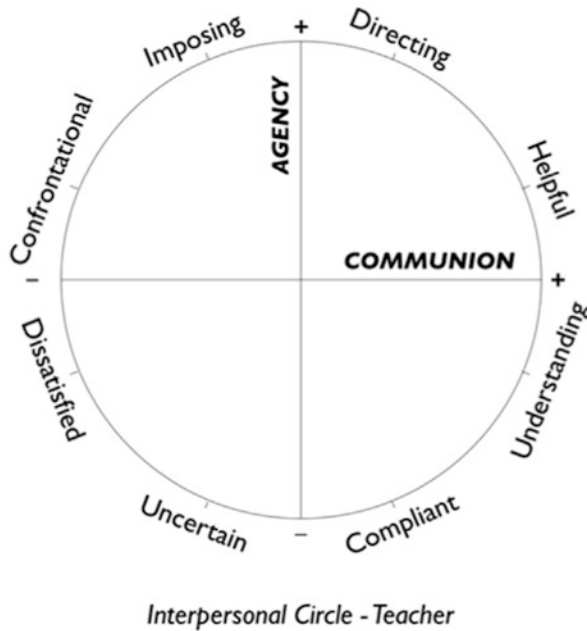


Fig. 22.1 The model for interpersonal teacher behavior (or teacher interpersonal circle). (Pennings et al., 2018)

empathic, social, harmonious or friendly (Gurtman, 2009). The agency dimension has also been referred to as the influence, control or power dimension of interpersonal relationships and the communion dimension as the proximity, warmth or affiliation dimension (Wubbels et al., 2012). Research on relationships and interactions between people in a variety of fields such as psychology, sociology, communication and even evolutionary biology has suggested that both of the two dimensions are at the same time necessary and sufficient to describe and analyse interpersonal relationships (Gaines et al., 1997; Leary, 1957; Lonner, 1980).

3.2 The Model of the Teacher Interpersonal Circle and Its Measurement

Within this chapter, we focus on studies investigating teacher-student relationships using the model of the Teacher Interpersonal Circle (Pennings et al., 2018). This model is an adaptation of more general models used in interpersonal theory (see also Leary, 1957) to the teacher-class relationship. It describes the teacher student relationship based on the agency and communion dimensions with eight interpersonal adjectives that represent various combinations of agency and communion

(see Fig. 22.1). Each adjective combines both dimensions and displays different degrees of agency and communion; for example, ‘directing’ teacher behaviour can be characterized as high on agency and moderate on communion, while ‘helpful’ behaviour is moderate on agency but high in terms of agency.

Studies investigating teacher-student interpersonal relationships have often focused on students’ perceptions of this behaviour and have measured these with the Questionnaire on Teacher Interaction (QTI; Wubbels et al., 2006). The QTI has eight scales corresponding with eight adjectives positioned around the interpersonal circle. Scales contain 3 to 12 items, depending on the version of the questionnaire used. There are versions of the QTI for different forms and types of education, such as primary, secondary and higher education, but also online education and supervisor-student interactions (Wubbels et al., 2014). It is a widely used instrument to measure perceptions of the teacher-student relationship. It has been used in more than 30 countries (Wubbels et al., 2006) and shown high construct validity and reliability (e.g., den Brok et al., 2006a). It also appears to be valid for measuring students’ perceptions of their teachers in various cultures (e.g., den Brok, et al., 2006b; den Brok & van Tartwijk, 2015).

While studies have shown that teacher-student interpersonal behaviours in the classroom can and do occur across the full interpersonal circumplex, healthy teacher-student interpersonal relationships have often been associated with high amounts of both teacher agency and communion. Teachers perceived by their classes as high on both agency and communion often have a relatively high sense of efficacy, a smaller chance for burnout, relatively high motivated students in their class, and are able to create learning environments that are both pleasant and safe, as well as varied and rich for learning (Wubbels et al., 2006). Interestingly, there are differences between teachers and students in associations between the two interpersonal dimensions and teacher versus student outcomes. For example, while for teacher well-being and positive emotions teacher agency is more predictive, for student outcomes teacher communion is more predictive (Donker et al., 2021). In the remainder of this chapter we zoom into the associations between teacher interpersonal agency and communion and student outcomes.

4 Teacher Agency, Communion and Student Outcomes

In this section we first discuss studies that have used the QTI and investigated associations with cognitive outcomes, such as achievement tests or report card grades. Subsequently, we discuss studies that have used the QTI and related teacher-student interpersonal behaviour to affective outcomes, such as subject-related attitudes and autonomous or intrinsic motivation. In doing so, we also indicate if covariates that were included in studies, such as prior outcomes, student characteristics or other context or learning environment characteristics.

4.1 *Student Achievement*

4.1.1 **Studies Using Dimensions of Interpersonal Relationships**

Studies using the QTI have been conducted in a variety of countries, ranging from Europe, Australia and the USA, to India and the Far East. When investigating associations between the two interpersonal dimensions and student achievement, studies mostly found positive associations of achievement with perceptions of both teacher agency and communion (e.g., Brekelmans, 1989; Georgiou & Kyriakides, 2012; Zijlstra et al., 2013). These associations were usually moderate to small. Effects were smaller in studies using multilevel analysis of variance and correcting for effects of student and teacher characteristics, than in studies investigating only the effect of interpersonal behaviours and not accounting for the hierarchical structure of collected data.

Zijlstra et al. (2013) reported that agency was a slightly stronger predictor for achievement than was communion. After control for prior achievement, about 5% of the differences in mathematics achievement in their study could be accounted for by both interpersonal dimensions. Interestingly, whereas the effect of agency on achievement appeared stable across classes, a differential effect could be found for communion. However, this differential effect could not be explained by variables such as class size, gender distribution, average class ability, teacher experience or the number of days a teacher taught the class per week. As their study was conducted in primary education, they argued that a potential effect for the stable findings for agency might lie in the lower self-regulatory skills of students, thus needing more agency by teachers.

In a study by Brekelmans (1989) on students' perceptions their relationship with their physics teachers in secondary education, perceptions on both dimensions were related to cognitive outcomes. The higher a teacher was perceived on the agency and communion dimension, the higher the outcomes of students on a physics test. In her study, teacher agency was the most important variable at the class level.

4.1.2 **Studies Using Sectors of Interpersonal Relationships**

Other studies did not investigate the association with the dimensions underlying the model, but instead focused on the associations with each of the scales (cf. Fig. 22.1). Positive correlations or regression coefficients were found for the directing scale and cognitive student outcomes (Goh & Fraser, 1998; Henderson & Fisher, 2008). In a study in Greece, Charalambous and Kokkinos (2018) also found positive associations between the directing scale and achievement in language and mathematics, as well as between supporting, understanding and compliant scales and achievement. However, they also found a negative association between the imposing scale and achievement in both school subjects, suggesting that teacher agency does not

always lead to high cognitive outcomes and that in the Greek context, communion may be more decisive than agency.

Strong and positive relationships with cognitive outcomes have also been found for the communion dimension and high communion related scales such as helpful and understanding (Goh & Fraser, 1998; Henderson & Fisher, 2008; Evans, 1998; see also Charalampous & Kokkinos, 2018). The more teachers were perceived as high on communion, the higher students' scores on cognitive tests. However, relationships between communion and cognitive outcomes were not always straightforward. In some studies, it could only be proven that low communion, or scores on the dissatisfied and confrontational scales, were related to lower performance, but not that scores on the helpful and understanding scales were related to higher performance (Rawnsley, 1997). In other studies, the relationship between communion and cognitive outcomes was not linear but curvilinear (i.e. lower perceptions of communion went together with low outcomes, but intermediate and higher values with higher performance until a certain ceiling of optimal communion was been reached; den Brok, 2001).

4.1.3 Other Findings Related to Student Achievement

Some studies found that only one of the two dimensions was related to student achievement, either agency (den Brok et al., 2004; Sivan & Chan, 2013) or communion (Bacete et al., 2014; Gupta & Fisher, 2008). A study by Gupta and Fisher (2008) reported a negative association of agency with student outcomes, where other studies reported mainly positive associations.

If report card grades were used as outcome measures, relationships with interpersonal behaviour were inconclusive (Levy et al., 1992; Telli et al., 2007). No relationship between student perceptions of communion and agency and their report card grades was found in these studies. A potential explanation might lie in that report card grades often are not just a measure of achievement, but are determined by other factors as well, such as affective factors and subjective factors, such as teacher expectations and beliefs (Brookhart et al., 2016).

When looking at the consistency of findings across contexts, higher associations have been found for both dimensions in mathematics and science than in (foreign) languages or social science classes (den Brok et al., 2004; Georgiou & Kyriakides, 2012). Within classes, different associations have been found for ethnic minority students and for mainstream students. den Brok et al. (2010) for example, found a positive association between teacher agency and report card grades for students with a Surinamese background in Dutch multicultural classes, but negative associations for students with parents born in the Netherlands and students with a Moroccan background, and no association for students with a Turkish background. In their study, no direct effects were found for communion on report card grades, but indirect effects were found for communion, with student motivation as a mediator.

4.2 *Affective Student Outcomes*

4.2.1 **Studies Using Dimensions of Interpersonal Relationships**

Studies using the two interpersonal dimensions all found a positive effect for both agency and communion on students' subject-related attitudes. Generally, effects of communion were stronger than those of agency.

For example, in a study of physics teachers and their students in the Netherlands, Brekelmans and her colleagues (Brekelmans, 1989; Brekelmans et al., 1990) found a stronger relationship between communion and students' attitudes than between agency and student attitudes: the stronger the perception of communion the more positive the attitude of the students towards the subject was. Also in a study of English as a foreign language (EFL) teachers in the Netherlands (den Brok et al., 2004) it was found that the effect of communion on students' pleasure in the subject was three to four times stronger than the effect of agency, even though both had a positive effect. For students' willingness to put effort in the subject and their degree of confidence in the subject, the association with communion was almost twice as large as the association with agency. In both studies the effects of agency and communion were corrected for the effect of student, class and teacher characteristics, such as gender, SES, class size, teacher gender, school type and report card grade. Moreover, these studies employed multilevel analysis techniques, thereby taking into account the effects of non-random sampling.

A study in Brunei (den Brok et al., 2005b) - also employing multilevel analyses and correcting the effect of interpersonal relationships for various student, class and teacher characteristics - indicated equally strong effects of agency and communion. However, that study was conducted with primary education science teachers and their students. A study on secondary science students and their teachers in India (den Brok et al., 2005a) again found similar positive associations of both agency and communion with students' attitudes towards science. In the study in India, multilevel analyses were conducted and associations were corrected for student covariates as well as other teaching variables.

A series of studies looking at both the dimensions of agency and communion in relation to affective outcomes in secondary school science was conducted in Turkey (den Brok et al., 2007; Telli et al., 2007, 2010). When looking at raw correlations, positive associations of agency were found with enjoyment of the subject, perceived usefulness of the subject, interest in the subject and time effort; however, correlations of communion with these variables was almost twice as high, except for effort where a similar correlation was found. In all cases, correlations were moderate to strong. Interestingly, after correcting for student, class and teacher covariates and conducting multilevel regression analyses, a less distinct pattern was found, showing small and positive associations between agency and enjoyment and interest, a small positive association of communion with interest, and no significant associations between the dimensions and the other outcome variables.

4.2.2 Studies Using Sectors of Interpersonal Relationships

Positive, strong associations have also been demonstrated between several QTI-scales, such as directing and helping, and subject-related attitudes, while negative relationships were found with the dissatisfied, confrontational, and, in most cases, the imposing scales (e.g., Evans, 1998; Goh & Fraser, 1998; Fisher et al., 1995; Henderson & Fisher, 2008; Rawnsley, 1997; van Amelsvoort, 1999). In most of these studies, all scales related significantly to student attitudes in terms of correlation coefficients – with directing, helpful, understanding and compliant relating positively; uncertain, dissatisfied, confrontational and imposing relating negatively – but only a small number of scales (e.g. supporting and understanding) remained statistically significant if the more conservative regression weights were used (e.g. den Brok et al., 2005b).

A number of these studies were conducted in Australia. Henderson and Fisher (2008), for example, studied Biology classes. In their study, they found that the QTI scales explained 33% of the variance in enjoyment, either uniquely or in combination with other learning environment variables. Evans (1998) studied Australian science classes and reported similar associations. Rawnsley (1997) studied mathematics teachers and again reported similar findings as in the other two mentioned Australian studies. Characteristic of these Australian studies is that they investigated the effects of interpersonal relationships taking into account other learning environment elements, but that respondent characteristics were not included. The studies indicated large amounts of variance explained jointly by interpersonal and other teacher behaviours (Rawnsley, 1997), while also a large amount of variance appeared to be explained by the QTI results uniquely.

In Greek classes, Charalampous and Kokkinos (2018) found positive correlations between scales displaying high communion and affective outcomes, such as attitudes towards language or mathematics and academic self-efficacy, while scales with low communion displayed negative correlations with these outcome variables.

Several studies investigating associations between QTI scales and attitudes have been conducted in Singapore, one with primary education mathematics classes (Goh & Fraser, 1998), one with secondary education science classes (Fisher et al., 1995), and two by Quek and her colleagues (Quek et al., 2005, 2007) in science classes. Interestingly, the authors of these studies report higher amounts of variance explained in student enjoyment than was the case in the Australian studies. Fisher et al. (1995), for example, reported a percentage of explained variance by interpersonal variables of 49%. This strong association was also reflected in correlation coefficients, ranging between $-.56$ (imposing) and $+.66$ (supporting). These patterns were similar in both studies. In a study on chemistry lessons (Quek et al., 2005), positive associations were reported for directing, helpful and understanding behaviour and negative associations were reported for uncertain, confronting and imposing. In that study, interpersonal variables explained twice as much variance in enjoyment as did other teaching or learning environment variables. In a study investigating attitudes to project work, Quek and her colleagues (Quek et al., 2007) reported a positive association between both the imposing and directing scales and

enjoyment (in project work), while a negative association was reported between imposing and attitude towards inquiry in project work. Overall, in their study low associations between teacher-student interpersonal relationships and affective outcomes were reported.

One other study was conducted in Korean science classes (Kim et al., 2000) and reported correlation coefficients ranging between $-.36$ (objecting) and $+.49$ (supporting). In all aforementioned studies, scales on the positive side of communion correlated positive, while scales on the negative side of communion correlated negatively.

In a study in Hong Kong, it was found that high communion scales displayed positive correlations with students' attitudes towards their teacher, their school subject as well as moral outcomes ($+.33$ to $+.71$), while low communion scales displayed negative associations with these variables ($-.25$ to $-.51$), with the imposing scale showing no correlation with these outcomes (Sivan & Chan, 2013).

In a study in Thailand, a negative association between the imposing scale and attitude towards English as a foreign language (EFL) was reported, but none of the other interpersonal scales was associated with attitude towards EFL (Wei & Onsawad, 2007).

4.2.3 Other Findings Related to Affective Outcomes

In an Indonesian study, associations were investigated between teacher agency and communion and student motivation in general, distinguishing between more autonomous forms and more controlled forms of motivation (Maulana et al., 2011). They found that both agency and communion were positively related to autonomous motivation and in similar strength, but that agency was more strongly related to controlled (or more extrinsic) motivation. They explained the latter finding by the cultural context of Indonesia, where high teacher agency in the classroom is both expected and valued.

A recent study in China investigated associations of teacher students' interpersonal relationships with student enjoyment and anxiety (Sun et al., 2018). It was found that only communion was moderately to strongly associated with these outcomes, being positively related to enjoyment and negatively to anxiety. However, the agency dimension was not significantly associated with either enjoyment or anxiety.

In a study by den Brok et al. (2010) in multicultural classes in the Netherlands, teacher-student communion showed strong associations with positive attitudes towards subject content among all cultural groups involved in their study. However, higher levels of teacher agency did not correlate with subject attitude among students with a Dutch background. For students with a Moroccan, Turkish or Surinamese background (but born in the Netherlands), higher levels of teacher agency had small to medium positive effects on subject attitude. The positive relationship between teacher agency and subject attitude might seem contrary to expectations based on the self-determination theory that predicts high motivation with

student autonomy and corresponding low teacher agency, but in recent applications of this theory to educational context, the importance of providing structure combined with autonomy is emphasized (Aelterman et al., 2018). Providing structure requires a certain level of teachers directiveness according to these authors. Another explanation might be that most multicultural schools in the Netherlands are situated in the major cities, where teaching is often rather challenging for teachers from a classroom management perspective (van Tartwijk et al., 2009). Low success in classroom management may result in low agency in student perceptions of the teacher-student relationship (Wubbels et al., 2006). Such low agency scores in these classes do not indicate high student autonomy, but rather disorder, which is negatively related with student motivations (Wubbels et al., 2006).

4.3 *Summary of Findings*

Overlooking all of the studies and their findings, some general trends could be seen. For achievement, both teacher agency and communion were positively related to student achievement, with the agency dimension (or its related scales) displaying stronger and more consistent associations with achievement than communion. For communion, associations were sometimes inconsistent or less straightforward. Associations of both dimensions or their related scales were more consistent for achievement tests than for report card grades.

For affective student outcomes, positive associations were also found with both teacher agency and communion, in this case communion showing stronger associations than agency. Findings showed some differences in strength depending on the type of affective outcome involved, but in all cases associations were positive.

As for both types of outcomes, it was found that associations of agency and communion often remained statistically significant if they were corrected for student or teacher covariates, as well as if they were combined with other teaching or learning environment variables. Also, while there was some variation between cultures, countries or school subjects, in general findings were consistent in the vast majority of studies.

5 Discussion

Research on teacher-student relationships has shown that warm and supportive relationships are positively related to students' affective learning outcomes, and via these outcomes - as well as directly - also to cognitive student outcomes (Cornelius-White, 2007; Roorda et al., 2011; Roorda et al., 2017; Witt et al., 2004). The present chapter reviewed research from an interpersonal (circumplex) theory perspective, including next to teacher warmth or interpersonal communion also a dimension depicting teacher authority or interpersonal agency.

Results of studies using the same instrument to link students' perceptions of the teacher-student relationship to student outcomes, namely the Questionnaire on Teacher Interaction (QTI) (Wubbels et al., 1985, 2006), showed an interesting picture. In most studies teacher agency positively, directly and in a stable and strong way related to student achievement. While communion related positively to achievement as well, this association was typically less strong than that of agency, and also less stable across classes, countries and contexts, and sometimes showed a more curvilinear association rather than a linear one. In this sense, the effect of communion on student achievement is complex: it may be that a minimum amount of communion is needed to enhance student achievement, but that too much communion may be detrimental, and that the optimal amount of communion to be supportive for achievement may be different for different students (Wubbels et al., 2023). The review did show that associations of both dimensions remained present after taking into account student, class or teacher background characteristics or other teaching or learning environment variables, although the effect would become smaller in most cases.

As for affective outcomes, most studies showed even stronger and positive associations with the two interpersonal dimensions of agency and communion than was the case for cognitive outcomes; in these cases, the association of communion was typically stronger than that of agency. These findings appeared rather consistently across countries, and remained as such after taking into account other covariates and learning environment variables. This finding may potentially be explained by the conditional nature of interpersonal relationships for the classroom climate and its effect on other teaching variables, which both directly and indirectly affect affective outcomes (Evertson & Weinstein, 2006; Fraser et al., 1987; van der Lans et al., 2020). Findings also appeared largely consistent for different affective variables, although effort and interest sometimes seemed to benefit slightly more from agency than did pleasure or autonomous motivation.

To some degree, the findings seem to confirm the potential intermediating role of affective outcomes in the relation between interpersonal relationships and cognitive outcomes (also see studies based on attachment theory and self-determination theory, Roorda et al., 2017). The intermediating effect can be inferred from the fact that stronger associations of the interpersonal relationship with affective outcomes were found than with cognitive outcomes; it suggests that both direct and indirect associations are at play, whereas the associations with cognitive outcomes are more direct. However, the findings also suggested that there is a direct relationship between the agency dimension of interpersonal relationship and cognitive outcomes, and that both dimensions of interpersonal relationships are relevant for student outcomes, separately as well as jointly. The present chapter did confirm prior findings that detrimental relationships can be characterized by opposition or conflict, but in addition showed that these relationships can also be typified by low agency, such as hesitancy.

Further research is needed to better understand what the precise interplay of both interpersonal dimensions is for student outcomes, what intermediate variables operate in this relationship, and if dimensions of the interpersonal relationship operate

more as conditional or as direct variables in their effect on student learning and outcomes. Combining insights from interpersonal and self-determination theory, where recently the role of structure for student motivation has been emphasized, might be useful when doing this. In this way, it can for example be investigated if structure in the classroom enhances (perceptions of) relations in the classroom, which in turn affect motivation, or if relations enhance the use of structure in the classroom, which in turn affect motivation. In general, research could further investigate the joint and unique effects and interplay of interpersonal relationships and other learning environment variables in relation to student outcomes, as we only understand the precise role of relationships on other environment variables to a limited degree (Fraser & Walberg, 2005). Also, since the dimensions may have different effects in different cultures or countries, more research is needed to understand what verbal and non-verbal behaviors play a role in this, and how moment-to-moment interactions determine the interpretation of relationships at the developmental level.

References

- Aelterman, N., Vansteenkiste, M., Haerens, L., Soenens, B., Fontaine, J. R. J., & Reeve, J. (2018). Toward an integrative and fine-grained insight in motivating and demotivating teaching styles: The merits of a circumplex approach. *Journal of Educational Psychology, 111*(3), 497–521.
- Bacete, F. J. G., Ferrá, P., Monjas, M. I., & Maranda, G. (2014). Teacher student relationships in first and second grade classrooms: Adaptation of the questionnaire on teacher interaction early primary. *Revista de Psicodidactica, 19*(1), 211–231.
- Brekelmans, M. (1989). *Interpersonal teacher behavior in the classroom*. W.C.C. [In Dutch].
- Brekelmans, M., Wubbels, T., & Créton, H. A. (1990). A study of student perceptions of physics teacher behavior. *Journal of Research in Science Teaching, 27*, 335–350.
- Brookhart, S., Guskey, T., Bowers, A. J., McMillan, J., Smith, L., Smith, J., & Welsh, M. (2016). A century of grading research: Meaning and value in the Most common educational measure. *Review of Educational Research, 86*(4), 803–848.
- Charalampous, K., & Kokkinos, C. M. (2018). The structure of pre-adolescents' perceptions of their teachers' interpersonal behaviours and their relation to pre-adolescents' learning outcomes. *Educational Studies, 44*(2), 167–189.
- Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A meta-analysis. *Review of Educational Research, 77*(1), 113–143.
- den Brok, P. (2001). *Teaching and student outcomes. A study on teachers' thoughts and actions from an interpersonal and a learning activities perspective*. W.C.C.
- den Brok, P. J., & van Tartwijk, J. W. F. (2015). Teacher-student interpersonal communication in international education. In M. Hayden, J. Levy, & J. Thompson (Eds.), *The Sage handbook of research on international education* (Vol. 2, pp. 309–324). Sage.
- den Brok, P., Brekelmans, M., & Wubbels, T. (2004). Interpersonal teacher behaviour and student outcomes. *School Effectiveness and School Improvement, 15*(3/4), 407–442.
- den Brok, P., Fisher, D., & Koul, R. (2005a). The importance of teacher interpersonal behaviour for secondary science students in Kashmir. *Journal of Classroom Interaction, 40*(2), 5–19.
- den Brok, P., Fisher, D., & Scott, R. (2005b). The importance of teacher interpersonal behaviour for student attitudes in Brunei primary science classes. *International Journal of Science Education, 27*(7), 765–779.

- den Brok, P., Brekelmans, M., & Wubbels, T. (2006a). Multilevel issues in studies using students' perceptions of learning environments: The case of the questionnaire on teacher interaction. *Learning Environments Research*, 9(3), 199–213.
- den Brok, P., Fisher, D., Brekelmans, M., Wubbels, T., & Rickards, T. (2006b). Secondary teachers' interpersonal behaviour in Singapore, Brunei and Australia: A cross-national comparison. *Asia-Pacific Journal of Education*, 26(1), 79–95.
- den Brok, P., Telli, S., & Cakiroglu, J. (2007). Teacher interpersonal behaviour and students' subject related attitudes in general and vocational science classes in Turkey. In C. S. Sunal & K. Mutua (Eds.), *Undertaking educational challenges in the 21st century: Research from the field* (Research on education in Africa, the Caribbean and the Middle East, Volume 5) (pp. 169–190). Information Age Publishing, Inc.
- den Brok, P., van Tartwijk, J., Wubbels, T., & Veldman, I. (2010). The differential effect of the teacher-student interpersonal relationship on student outcomes for students with different ethnic backgrounds. *British Journal of Educational Psychology*, 80, 199–221.
- Donker, M. H., van Vemde, L., Hessen, D. J., van Gog, T., & Mainhard, T. (2021). Observational, student, and teacher perspectives on interpersonal teacher behavior: Shared and unique associations with teacher and student emotions. *Learning and Instruction*, 73, 101414.
- Evans, H. (1998). *A study on students' cultural background and teacher-student interpersonal behaviour in secondary science classrooms in Australia*. Unpublished doctoral dissertation, Curtin University of Technology.
- Evertson, C. M., & Weinstein, C. S. (2006). Classroom management as a field of inquiry. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 3–15). Erlbaum.
- Fisher, D., Henderson, D., & Fraser, B. (1995). Interpersonal behaviour in senior high school biology classes. *Research in Science Education*, 25(2), 125–133.
- Fraser, B. J. (2014). Classroom learning environments: Historical and contemporary perspectives. In N. Lederman & S. Abell (Eds.), *Handbook of research on science education volume II* (pp. 104–119). Routledge.
- Fraser, B. J., & Walberg, H. J. (2005). Research on interpersonal relationships and learning environments: Context, retrospect and prospect. *International Journal of Educational Research*, 43(1–2), 103–109.
- Fraser, B. J., Walberg, H. J., Welch, W. W., & Hattie, J. A. (1987). Syntheses of educational productivity research. *International Journal of Educational Research*, 11, 145–252.
- Gaines, S. O., Panter, A. T., Lyde, M. D., Steers, W. N., Rusbult, C. E., Cox, C. L., & Wexler, M. O. (1997). Evaluating the circumplexity of interpersonal traits and the manifestation of interpersonal traits in interpersonal trust. *Journal of Personality and Social Psychology*, 73, 610–623.
- Georgiou, M., & Kyriakides, L. (2012). The impact of teacher and principal interpersonal behavior on student learning outcomes: A large scale study in secondary schools of Cyprus. In T. Wubbels, P. den Brok, J. van Tartwijk, & J. Levy (Eds.), *Interpersonal relationships in education: An overview of contemporary research* (pp. 119–136). Sense Publishers.
- Goh, S. C., & Fraser, B. J. (1998). Teacher interpersonal behaviour, classroom environment and student outcomes in primary mathematics in Singapore. *Learning Environments Research*, 1, 199–229.
- Granic, I., & Hollenstein, T. (2003). Dynamic systems methods for models of developmental psychopathology. *Development and Psychopathology*, 15(3), 641–669.
- Gupta, A., & Fisher, D. (2008). Teacher-student interactions in a technology-supported science classroom environment in relation to selected learner outcomes: An Indian study. In *Proceedings international conference on science*. Mathematics and Technology Education.
- Gurtman, M. B. (2009). Exploring personality with the interpersonal circumplex. *Social Psychology Compass*, 3, 1–19.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Henderson, D. G., & Fisher, D. L. (2008). Interpersonal behaviour and student outcomes in vocational education classes. *Learning Environments Research*, 11, 19–29.

- Hughes, J. N., Cavell, T. A., & Willson, V. (2001). Further support for the developmental significance of the quality of the teacher–student relationship. *Journal of School Psychology, 39*, 289–301.
- Kim, H. B., Fisher, D. L., & Fraser, B. J. (2000). Classroom environment and teacher interpersonal behaviour in secondary science classes in Korea. *Evaluation and Research in Education, 14*(1), 3–22.
- Leary, T. (1957). *An interpersonal diagnosis of personality*. Ronald Press Company.
- Levy, J., Wubbels, T., & Brekelmans, M. (1992). Student and teacher characteristics and perceptions of teacher communication style. *Journal of Classroom Interaction, 27*, 23–29.
- Lonner, W. J. (1980). The search for psychological universals. In H. C. Triandis & W. W. Lambert (Eds.), *Handbook of cross-cultural psychology* (pp. 143–204). Allyn & Bacon.
- Martin, A. J., & Dawson, M. (2009). Interpersonal relationships, motivation, engagement, and achievement: Yields for theory, current issues, and educational practice. *Review of Educational Research, 79*(1), 327–365.
- Maulana, R., Opdenakker, M. C., den Brok, P., & Bosker, R. (2011). Teacher-student interpersonal relationships in Indonesia: Profiles and importance to student motivation. *Asia Pacific Journal of Education, 31*(1), 33–49.
- Mehrabian, A. (1981). *Silent messages*. Wadsworth Publishing Company.
- Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art – Teacher effectiveness and professional learning. *School Effectiveness and School Improvement, 25*(2), 231–256.
- Pace, J. L., & Hemmings, A. (2007). Understanding authority in classrooms: A review of theory, ideology and research. *Review of Educational Research, 77*(1), 4–27.
- Pennings, H. J. M., Brekelmans, M., Sadler, P., Claessens, L. C. A., van der Want, A. C., & van Tartwijk, J. (2018). Interpersonal adaptation in teacher-student interaction. *Learning and Instruction, 55*, 41–57.
- Quek, C. L., Wong, A. F. L., & Fraser, B. J. (2005). Student perceptions of chemistry laboratory environments, student-teacher interactions and attitudes in secondary school gifted education classes in Singapore. *Research in Science Education, 35*, 299–321.
- Quek, C. L., Wong, A. F. L., Divarahan, S., Liu, W. C., Peer, J., & Williams, M. D. (2007). Secondary school students' perceptions of teacher-student interaction and students' attitudes towards project work. *Learning Environments Research, 10*, 177–187.
- Rawsley, D. G. (1997). *Associations between classroom learning environments, teacher interpersonal behaviour and student outcomes in secondary mathematics classrooms*. Unpublished doctoral dissertation, Curtin University of Technology.
- Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher-student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research, 81*(4), 493–529.
- Roorda, D. L., Jak, S., Zee, M., Oort, F. J., & Koomen, H. M. Y. (2017). Affective teacher–student relationships and students' engagement and achievement: A meta-analytic update and test of the mediating role of engagement. *School Psychology Review, 46*(3), 239–261.
- Ryan, R. M., & Deci, E. L. (2000). Self determination theory and the facilitation of intrinsic motivation, social development and well-being. *American Psychologist, 55*(1), 66–78.
- Sadler, P., Ethier, N., & Woody, E. (2011). Interpersonal complementarity. In L. M. Horowitz & S. Strack (Eds.), *Handbook of interpersonal psychology* (pp. 123–142). Wiley.
- Schrod, P., Witt, P. L., Myers, S. A., Turman, P. D., Barton, M. H., & Jernberg, K. A. (2008). Learner empowerment and teacher evaluations as functions of teacher power use in the college classroom. *Communication Education, 57*(2), 180–200.
- Sivan, A., & Chan, D. W. K. (2013). Teacher interpersonal behavior and secondary students' cognitive, affective and moral outcomes in Hong Kong. *Learning Environments Research, 16*(1), 23–36.
- Strack, S., & Horowitz, L. M. (2011). Introduction. In L. M. Horowitz & S. Strack (Eds.), *Handbook of interpersonal psychology: Theory, research, assessment, and therapeutic interventions* (pp. 1–13). Wiley.

- Sun, X., Mainhard, T., & Wubbels, T. (2018). Development and evaluation of a Chinese version of the Questionnaire on Teacher Interaction (QTI). *Learning Environments Research*, 21(1), 1–17.
- Telli, S., den Brok, P., & Cakiroglu, J. (2007). Students' perceptions of science teachers' interpersonal behaviour in secondary schools: Development of the Turkish version of the questionnaire on teacher interaction. *Learning Environments Research*, 10(2), 115–129.
- Telli, S., Den Brok, P., & Cakiroglu, J. (2010). The importance of teacher-student interpersonal relationships for students' subject-related attitudes in Turkey. *Research in Science and Technological Education*, 28(3), 261–276.
- Thijs, J. T., & Koomen, H. M. Y. (2008). Task-related interactions between kindergarten children and their teachers: The role of emotional security. *Infant and Child Development*, 17, 181–197.
- van Amelsvoort, J. (1999). *Perspective on instruction, motivation and self-regulation* [In Dutch]. Unpublished doctoral dissertation, Katholieke Universiteit Nijmegen.
- van der Lans, R., Cremers, J., Klugkist, I., & Zwart, R. (2020). Teachers' interpersonal relationships and instructional expertise: How are they related? *Studies in Educational Evaluation*, 66, 100902.
- van Tartwijk, J. (1993). *The interpersonal significance of nonverbal behaviour in the classroom* (in Dutch: De interpersoonlijke betekenis van nonverbaal gedrag van docenten in de klas). WCC.
- van Tartwijk, J., den Brok, P., Veldman, I., & Wubbels, T. (2009). Teachers' practical knowledge about classroom management in multicultural classrooms. *Teaching and Teacher Education*, 25(3), 453–460.
- Watzlawick, P., Beavin, J. H., & Jackson, D. (1967). *The pragmatics of human communication*. Norton.
- Wei, M., & Onswad, A. (2007). English teachers' actual and ideal interpersonal behaviour and students' outcomes in secondary schools of Thailand. *The Journal of Asia TEFL*, 4(2), 1–29.
- Witt, P. L., Wheelless, L. R., & Allen, M. (2004). A meta-analytical review of the relationship between teacher immediacy and student learning. *Communication Monographs*, 71(2), 184–207.
- Woolfolk Hoy, A., & Weinstein, C. S. (2006). Student and teacher perspectives on classroom management. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice, contemporary issues* (pp. 181–223). Erlbaum.
- Wubbels, T., Créton, H. A., & Hooymayers, H. P. (1985). Discipline problems of beginning teachers, interactional teacher behaviour mapped out. *Abstracted in Resources in Education*, 20(12), 153. (ERIC Document Reproduction Service No. 260040).
- Wubbels, T., Créton, H., & Holvast, A. J. C. D. (1988). Undesirable classroom situations. *Interchange*, 19, 25–40.
- Wubbels, T., Brekelmans, M., den Brok, P., & van Tartwijk, J. (2006). An interpersonal perspective on classroom management in secondary classrooms in The Netherlands. In C. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice and contemporary issues* (pp. 1161–1191). Lawrence Erlbaum Associates.
- Wubbels, T., Brekelmans, J. M. G., den Brok, P. J., van Tartwijk, J. W. F., Mainhard, T., & Levy, J. (2012). Let's make things better! Developments in research on interpersonal relationships in education. In T. Wubbels, P. J. den Brok, J. van Tartwijk, & J. Levy (Eds.), *Interpersonal relationships in education: An overview of contemporary research* (pp. 225–249). SENSE Publishers.
- Wubbels, T., Brekelmans, M., den Brok, P., Wijsman, L., Mainhard, T., & van Tartwijk, J. (2014). Teacher-student relationships and classroom management. In E. T. Emmer, E. Sarbonie, C. Evertson, & C. Weinstein (Eds.), *Handbook of classroom management* (2nd ed., pp. 363–386). Taylor and Francis Publishers.
- Wubbels, T., Mainhard, T., den Brok, P., Claessens, L., & van Tartwijk, J. (2023). Classroom management at different timescales: An interpersonal perspective. Accepted for publication. In E. Sabornie & D. Espelage (Eds.), *Handbook of research on classroom management: Research, practice and issues* (3rd ed.) (pp. 388–414). New York: Routledge.
- Zijlstra, Z., Wubbels, T., Brekelmans, M., & Koomen, H. M. Y. (2013). Teacher interpersonal behavior and associations with mathematics achievement in Dutch early grade classrooms. *The Elementary School Journal*, 113(4), 517–540.

Perry den Brok is a professor of learning and education sciences at Wageningen University and Research, The Netherlands. His research focuses on all types of education, and particularly on topics such as learning environments, teacher behaviour, teacher professional learning and development, and educational innovation. He is also chair of the 4TU Centre for Engineering Education, a centre focusing on innovation in higher education.

Jan van Tartwijk is a professor of education and director of the Graduate School for Teaching, Utrecht University in Utrecht, The Netherlands. In his research, he focuses on teacher student communication processes, learning and assessing learning at the workplace, teacher education and expertise development.

Tim Mainhard is professor in educational sciences at the Institute of Education and Child Studies at Leiden University. His research focusses on social dynamics in educational settings and their impact on student and teacher outcomes. Tim teaches in the teacher education programme of Leiden University.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.



The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 23

Exploring How Teachers' Personal Characteristics, Teaching Behaviors and Contextual Factors Are Related to Differentiated Instruction in the Classroom: A Cross-National Perspective



Annemieke Smale-Jacobse, Peter Moorer, Ridwan Maulana , Michelle Helms-Lorenz , Carmen-María Fernández-García, Mercedes Inda-Caro, Seyeoung Chun, Abid Shahzad, Okhwa Lee, Amarjargal Adiyasuren, Yulia Iridayanti, Ulziisaikhan Galindev, and Nurul Fadhilah

Abstract Internationally, differentiated instruction (DI) is suggested as a teaching approach that can help teachers to meet the varying learning needs of students in the classroom. However, not all teachers reach a high level of implementation. Personal characteristics of the teacher as well as teaching quality may affect the degree and quality of DI. In addition, several classroom-, school-, and country characteristics may affect DI practices. In this chapter, literature is reviewed about personal factors, teaching characteristics and contextual factors influencing DI. Findings from the literature are connected to analyses of classroom observation-data collected in six

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-31678-4_23.

A. Smale-Jacobse (✉)
Department of Teacher Education, University of Groningen, Groningen, the Netherlands
Hanze University of Applied Sciences, Groningen, the Netherlands
e-mail: a.e.smale@pl.hanze.nl

P. Moorer · R. Maulana · M. Helms-Lorenz
Department of Teacher Education, University of Groningen, Groningen, the Netherlands

C.-M. Fernández-García · M. Inda-Caro
Department of Educational Sciences, University of Oviedo, Oviedo, Spain
e-mail: fernandezcarmen@uniovi.es; indamaria@uniovi.es

S. Chun
Department of Education, Chungnam National University, Daejeon, South Korea

A. Shahzad
Department of Education, The Islamia University, Bahawalpur, Pakistan

countries including Indonesia, the Netherlands, Mongolia, Pakistan, South Korea and Spain. The chapter aims to contribute to insights into factors related to DI and into differences in these associations between the six countries. This chapter concludes by discussing scientific and practical implications.

1 Introduction

Globally, teachers are challenged to meet the learning needs of groups of students with heterogeneous characteristics. Students may, for instance, vary in their readiness, interests and learning preferences (Tomlinson et al., 2003). Heterogeneity in classrooms is becoming larger with increasing inclusion of students with disabilities, different backgrounds and varying experiences into contemporary classrooms around the world (Rock et al., 2008; UNESCO, 2017, 2020a). As suggested in several theoretical frameworks, such as Vygotsky's zone of proximal development (Vygotskii & Cole, 1978), self-determination theory (Deci & Ryan, 1985) and the theory of flow (Csikszentmihályi, 2008), learning occurs best when instruction matches students' needs. Internationally, the question of how to deal with varying learning needs is currently approached by suggesting inclusive educational systems in which differentiated instruction (DI) or other types of adaptive instruction are used to match instruction to students' needs (UNESCO, 2017, 2020a). DI is defined as the adaptation of content, process, product, learning environment or learning time based on information about students' readiness or another relevant student characteristic (such as learning preference or interest) with the goal to better align teaching to students' needs (Smale-Jacobse et al., 2019). Teachers using DI proactively offer different 'routes' in their lessons for students to reach the learning goals. By doing so, the learning can be better adjusted to students' needs. DI has been a much-studied topic across various countries (Sun & Xiao, 2021). Multiple studies

O. Lee

Department of Education, Chungbuk National University, Cheongju, South Korea

A. Adiyasuren

Mongolian National University of Education, Ulaanbaatar, Mongolia

Y. Irnidayanti

Department of Biology and Biology Education, Faculty of Mathematics and Science, State University of Jakarta, Jakarta, Indonesia

U. Galindev

The Department of Educational Administration, Mongolian National University of Education, Ulaanbaatar, Mongolia

e-mail: ulziisaikhan@msue.edu.mn

N. Fadhilah

Department of Biostatistics and Population Studies, Faculty of Public Health, Universitas Indonesia, Depok, Indonesia

have shown that DI can lead to better learning outcomes, although more evidence about the effectiveness of different applications of DI is still needed (Deunk et al., 2018; Smale-Jacobse et al., 2019; Steenbergen-Hu et al., 2016).

Although DI seems to entail useful pedagogical-didactical approaches for student-centered teaching, implementation can be challenging. In general, teachers acknowledge the need to address students' varying needs, but they typically show little differentiation in their lessons (Tomlinson et al., 2003). Factors like the knowledge or skills of a teacher may affect the implementation of DI, besides the impact of contextual factors like the school system or cultural beliefs in society (Loreman et al., 2014; UNESCO, 2020a). A recent narrative review of studies from different countries showed that contextual factors like class size, time constraints and density of the curriculum were related to the implementation of DI, as well as personal characteristics of the teacher (Lavania & Nor, 2020). Thus, when aiming to gain insight into how implementation of DI may be improved, we should take into account factors regarding the context in which DI is executed and characteristics of the teacher that may influence implementation. Research into contextual factors that influence the implementation of DI is relatively scarce up to date (Sun & Xiao, 2021). Since factors related to the teacher and the context may vary across educational systems and countries, studying these influences with international data can give valuable insights in similarities and differences across countries.

Helms-Lorenz and Visscher (2021) identified different relevant contextual factors influencing teaching behavior including class size, student performance in the class, school policy, leadership and educational policies of the country. In the same vein, Brühwiler and Blatchford (2011) summarized several factors influencing teachers' adaptive instruction and, eventually, student performance in a theoretical model. At the teacher level, the authors included general characteristics like gender, teaching experience, personal motivation, affect and competency that may influence teaching. Furthermore, variables referring to the context of the classroom like class size and heterogeneity of the classroom are hypothesized to influence adaptive teaching. At a higher level, factors like characteristics of the educational system of the country or region are mentioned. As identified in the dynamic model of teaching (Kyriakides et al., 2009), national and regional educational policy influences school policy, which in turn may affect teaching.

In this chapter, we aim to explore the relationship between the implementation of DI and various personal characteristics, teaching behaviors and contextual factors. We will study this using empirical data from secondary schools in six different countries to explore the relations across a rich set of different contexts. First, let us turn to the literature about the influence of variables included in the study. In line with the model of Brühwiler and Blatchford (2011), we will discuss findings from the literature across different categories: classroom (teaching) processes, teacher characteristics, classroom context, school context and country (educational system).

1.1 Classroom Processes

1.1.1 Differentiated Instruction

Across educational contexts, policy makers and teachers stressed the need to use frequent assessment and to adapt the curriculum towards individual learning needs (OECD, 2012; UNESCO, 2020a; UNESCO, 2017). Yet, observational studies in secondary education found that teachers across different countries in general did not show much DI in their lessons (Maulana et al., 2021; Van der Lans et al., 2017). Nevertheless, teachers' DI can develop in contexts in which DI is explicitly promoted (Bondie et al., 2019; Schipper et al., 2017). In literature on teaching and teaching effectiveness, DI is recognized as one of the key characteristics of effective teaching (Kyriakides et al., 2009; Seidel & Shavelson, 2007; de Grift & Wim, 2014).

1.1.2 Differentiated Instruction and Other Effective Teaching Behaviors

Most models of DI stress the interrelatedness of DI and other teaching behaviors. For instance, in the differentiation model of Tomlinson (2014), DI is said to be influenced by general principles of differentiation like high-quality curriculum, teaching up and continuous assessment. In addition, teaching behaviors like stimulating mutual respect and supporting students to have high expectations of what they can do are important factors that may help set the stage for DI (Tomlinson & Imbeau, 2010). In the description of DI principles by Van Geel et al. (*this book*), general teaching quality indicators like communicating clear lesson goals, introducing the lesson and monitoring students' progress have a central place. The same goes for the model of Smale-Jacobse et al. (2019) in which DI is embedded in a context of continuous assessment, high-quality teaching and curriculum and a supportive learning environment. In that sense, other teaching behaviors are hypothesized to be related to teachers' DI. In some models of teaching quality, differentiation is viewed as a high-quality dimension of general teaching quality indicators like questioning, modeling or assessment (Kyriakides et al., 2009). Observational studies showed that teachers who have highly developed basic teaching skills are typically more likely to differentiate (Van der Lans et al., 2017). DI has often been found to be one of the more complex domains of teaching, clustering together with other complex teaching skills like activating students and teaching learning strategies (Van der Lans et al., 2017). In our study, DI is conceptualized as one of six domains of effective teaching behavior: creating a safe learning climate, efficient classroom management, quality of instruction, activating teaching methods, teaching learning strategies and differentiated instruction (de Grift & Wim, 2014). Interrelatedness between DI and teaching behaviors in other domains was previously found in all of the countries included in the current empirical study (Chun et al., 2020; Maulana et al., 2021).

1.1.3 Other Classroom Processes

Besides the teaching behaviors described above, which were included in our study, there are other classroom processes that may be related to DI. One example is the interpersonal relationship between the teacher and the students. A previous study shows that students rather uniformly perceive teachers who show relatively high-quality DI to be “helpful” or “directing” in their interactions (Van der Lans et al., 2020).

1.2 Teacher Characteristics

1.2.1 Teaching Experience

A personal factor of the teacher that may affect the implementation of DI is teaching experience. Beginning teachers are often still developing basic teaching skills and are generally relatively inflexible in their teaching. Experienced teachers, on the other hand, are generally better at offering challenging curricula, they often have deep representations of the subject matter and are skilled in monitoring and providing feedback (Berliner, 2004). Expert teachers often have a broad pedagogical and didactical repertoire and are typically more able to evaluate students' learning needs (Hayden et al., 2013). This could make it easier for them to flexibly adapt their teaching to students' needs. Fullers' (Fuller, 1969) theory of teacher development posits that teachers typically shift their concern from a focus on themselves to a focus on the task and later on to a focus on the impact of their teaching for students. Secondary school teachers generally experience a shift in focus during their careers, developing from an emphasis on the subject matter to an emphasis on gaining didactical and pedagogical expertise (Beijaard et al., 2000). The latter, more student-centered focus in both theories of teacher development seems to be more in line with the student-centered philosophy of DI.

Teaching experience was found to be positively related to DI in the Netherlands (Van der Pers & Helms-Lorenz, 2019), Indonesia (Suprayogi et al., 2017), and in countries not included in our study like Singapore and the United States (Van Tassel-Baska et al., 2008). However, there are also studies in which less-experienced teachers differentiated better than more-experienced counterparts, for instance in Spain and South Africa (Fernández-García et al., 2019; De Jager et al., 2017). In Spain, the current teacher-training program includes increased attention for pedagogical, didactical and psychological aspects of working with students, which may explain why novice teachers show higher quality DI in this county (Fernández-García et al., 2019). In Mongolia, about half of all teachers have between 1–10 years of experience (Ministry of Education and Science, 2021). In Pakistan, teachers on average have about 7 years of experience with a maximum of around 30 years. In South Korea, teachers in lower secondary education on average have around 16 years of

experience. Since about one third of all teachers are 50 years or older, many new teachers will be starting in the coming years though (OECD, 2019b). Differences in the relations between experience and DI may be caused by variation in the way teachers are prepared for DI in teacher education, in-service professionalization or by differences in educational policy (De Neve & Devos, 2016; De Jager et al., 2017), which stresses the need to take the broader context into account.

1.2.2 Teacher Gender

Teacher gender might be a less obvious influence on DI than experience. However, since there are studies pointing at gender differences in teaching styles, teacher gender is a characteristic worth exploring. In most of the countries included in our sample, there are both female and male teachers in secondary education. In Pakistan and South Korea, there are relatively more female teachers in lower secondary education. In the Netherlands, Spain and Indonesia the proportion of female and male teachers in secondary education is relatively equal (UNESCO, 2021). Alternatively, in Mongolia, more than 80% of all secondary school teachers are female (Ministry of Education and Science, 2021).

When turning to the relations between gender and teaching, there are some studies pointing at advantages for female teachers. For instance, a study using student-ratings found that Spanish female teachers in secondary and vocational education were rated higher than male teachers regarding their implementation of DI and several other domains of teaching (Fernández-García et al., 2019). In the same vein, an observational study executed in the Netherlands found female pre-service teachers to ensure a better learning climate and have better quality of instruction (Maulana & Helms-Lorenz, 2017).

However, there are also studies in which male teachers seemed to have an advantage over female teachers or in which there were little gender effects on teaching quality. In a study in Flanders, for instance, male teachers evaluated themselves more positively on leadership qualities and on helpful/friendly interpersonal behavior (Van Petegem et al., 2005). A study in the Netherlands showed that students evaluated male teachers as more cooperative and friendly than female teachers (Opdenakker et al., 2012). Another study found gender effects in favor of males in teaching learning strategies (Van der Pers & Helms-Lorenz, 2019).

It seems that gender differences in teaching are mixed depending on the context, the measurement instrument and the teaching domains. Findings in favor of males were found regarding classroom management and interpersonal relationships with students. One study executed in Spain reported that females were better in DI (Fernández-García et al., 2019), but other studies did not report on direct relations between gender and DI.

1.2.3 School Subject

There are studies arguing that the way a school subject is perceived by teachers can influence their teaching (Grossman & Stodolsky, 1995). In the countries included in our sample, many different school subjects are taught ranging from about 8–20 core subjects followed by students. Turning to between-subject differences in DI, prior studies did not find evidence for pronounced differences. In a study of Pozas et al. (2020) in which teachers were questioned about their DI, a rather similar response pattern was found for both German and Mathematics. There were slight differences though, with mathematics teachers using (peer)tutoring more and German teachers indicating more use of project-based learning. In a study in which lessons of pre-service teachers in the Netherlands were observed, no significant differences in teaching quality were found across school subjects (Maulana & Helms-Lorenz, 2017).

1.2.4 Other General Characteristics of the Teacher

In addition to the previously mentioned teacher characteristics included in our study, there are other teacher characteristics that could be related to DI. In prior studies, characteristics of teachers like knowledge, growth mindset, beliefs, self-efficacy and professional vision were related to the implementation of DI (Coubergs et al., 2017; Suprayogi et al., 2017; Vantieghem et al., 2020; UNESCO, 2020a; Whitley et al., 2019). There are between-country differences that may affect such teacher characteristics. For instance, in South Korea only top students from high schools can enter teacher-training programs, which makes for highly knowledgeable and skilled teacher-candidates. Conversely, while in countries like Indonesia, Pakistan and Mongolia teaching is a relatively low-paid profession that does not attract many of the top graduates. In addition, the curricula of the teacher training programs and the professionalization initiatives may affect teachers' knowledge, skills and beliefs. There are differences between countries with respect to how well teachers feel prepared for pedagogical and didactical issues in classroom practice. For instance, in Spain and the Netherlands, only about a quarter of all teachers reported to feel prepared to teach in mixed-ability classrooms (OECD, 2019b). In Mongolia, there is increasing attention for teacher training and professionalization, but up to date a wide variety of approaches is used across the country (UNESCO, 2020b). And teacher training programs in Pakistan and Indonesia are not yet up to international standards (United States Agency for International Development, 2006; World Bank, 2015). From the countries included in our sample, teachers are particularly valued and supported in South Korea (OECD, 2016a).

1.3 Classroom Context

1.3.1 Class Size

The majority of studies on class size have reported that within smaller classes, teachers have more care for students' individual needs than in larger classes. Blatchford et al. (2011) found that students in smaller classes received more attention and had more active interactions with the teacher. Another study reported that teachers in smaller classes devoted less time to group instruction and more time to individual instruction, especially in below-average classes (Betts & Shkolnik, 1999). Observational studies in Dutch secondary education showed that, on average, teachers use DI more in smaller classes (Maulana & Helms-Lorenz, 2017; Van der Pers & Helms-Lorenz, 2019). Teachers typically perceive it as a relatively time-demanding and difficult to adapt their instruction to the substantial spread of learning needs in large classes (Roiha, 2014; Wan, 2014). Across OECD countries and economies, teachers who teach larger classes tend to spend less classroom time on actual teaching and learning (OECD, 2019b).

Although overall findings point in the direction of DI being easier for teachers to implement in smaller classrooms, the link between the two is not always clear. For instance, in the study of Suprayogi et al. (2017), Indonesian teachers reported slightly more DI in larger classes. In the study of Brühweiler and Blatchford et al. (2011), class size was not directly related to classroom processes nor student outcomes in secondary education. This illustrates that, although smaller classes may make DI easier, lower class size does not by definition affect teaching nor student outcomes. In fact, teaching quality has been suggested to impact students more than class size (OECD, 2010).

In the countries included in our sample, the average class size differs considerably. In countries like Mongolia, Spain, South Korea and the Netherlands, the average class size is around the OECD average of 21 students (Education policy and data center, 2018; OECD, 2021). In the Netherlands, class size differs substantially between different educational tracks (Van Bergen et al., 2016). In Mongolia, class size differs considerably from around 15 students per teacher in rural areas up to 60 students per teacher in urban areas (UNESCO, 2019). The average class size in Pakistan is typically large, more than 40 students per class is not exceptional. In Indonesia, class size is also relatively large, with estimates of average class size ranging from about 33 to 47 students per teacher (Hendayana et al., 2010; OECD, 2014a).

1.3.2 Other Classroom Context Factors

Besides class size, another factor that may be related to the implementation of DI is the heterogeneity of the classroom. A large spread of learning needs can make it challenging for teachers to cater to individual students (Wan, 2014). On the other hand, external differentiation between classes may impede differentiation practices within the classroom. For instance, in Dutch secondary education students are tracked early on based on (presumed) abilities. Therefore, secondary school

teachers generally feel less need for DI than in primary education (Van Casteren et al., 2017), although there is in fact still large variation in attainment within the tracks (OECD, 2016b). In most countries in our sample, students first follow compulsory lower secondary education in mixed ability classes for 2 to 4 years. This could imply that classes in these countries are relatively more heterogeneous than in Dutch lower secondary education. Nevertheless, about half of all Dutch teachers do report to have more than 10% of students with special needs in their classes, illustrating that there may be other sources of heterogeneity too (OECD, 2019b). In upper secondary education, students are split up across different ability tracks varying from two different levels – an academic track and a vocational/technical track – in Spain, to six different ability tracks in the Netherlands (early tracking). Alternatively, in Mongolia and Indonesia, most students stay in their heterogeneous classes in upper secondary education. However, there are also students that switch to a different institution for vocational/technical education. In Pakistan, students choose between general and technical/vocational education before entering secondary education. After that, students are not split up further based on their abilities either but they do choose between different electives. In South Korea, upper secondary students can enroll in various types of high schools like general high schools, vocational high schools, science high schools or special high schools.

A teacher may additionally let the SES or the cognitive composition of the class influence the way they choose to implement DI, for instance by taking into account that homogeneous grouping could be detrimental for low-achieving students (Deunk et al., 2018). In addition, the cultural composition of a class may drive teachers towards differentiated approaches aimed at culturally responsive teaching (Gay, 2013). In Spain, for instance, an above-average percentage of students is born in another country (OECD, 2016c), which may make classes more culturally diverse.

1.4 School Context

Although the effects of school factors on instructional quality are typically small (Opdenakker & Van Damme, 2007), there are ways that schools can support, or hinder, teachers in their implementation of DI. Several aspects of the school climate may influence teaching and learning. School climate includes school organization, relations in the school community, leadership, available resources and institutional and structural features of the school environment to name a few (Wang & Degol, 2016). In the Netherlands and South Korea, schools have much autonomy over their resources and curriculum, while schools in Spain have somewhat less autonomy (OECD, 2011). In Mongolia, schools have little autonomy in matters of resources or curriculum. Also, in Indonesia and Pakistan, a standardized curriculum determined by the government is followed.

Several studies show that school principals can play an important role in teachers' willingness and ability to differentiate instruction (Goddard et al., 2010; Hertberg-Davis & Brighton, 2006). At the school level, working together with colleagues in a 'pedagogical team culture' may enhance teachers' implementation of

DI (Smit & Humpert, 2012). Additionally, the way schools are set up may influence DI. For instance, schools may vary in flexibility to move between different tracks (Gamoran, 1992). Moreover, school-level practices like providing enough preparation time for teachers may affect DI. Various studies show that teachers often experience lack of time for preparation and implementation of DI (De Jager, 2017; De Jager, 2013; Lavania & Nor, 2020; Roiha, 2014).

1.5 *Characteristics of the Country*

Based on a large-scale study on teaching quality across European, North-American, Pacific Countries, Canada and Australia, Reynolds et al. (2002) concluded that most factors known from national school- and teacher effectiveness research ‘work’ in different international contexts. However, there are country-specific differences in how teaching behaviors are interpreted and valued. The six countries included in the current study differ in many ways, for instance in the way education is organized, how the teaching profession is set up and valued, and what the classroom context is like. Some specifics of these countries that could affect DI through classroom processes, characteristics of the teachers, and the context of the school have been discussed above. In this paragraph, we will discuss some general country characteristics, policies related to DI and country-specific resources.

International comparisons of student performance show that students from South Korea are among the top performers internationally. Dutch students show above average performance in comparison to other countries and the performance of Spanish students is around the OECD average in the PISA evaluation. Indonesia is positioned among the lowest performing educational systems (Mullis et al., 2020; Mullis et al., 2017; OECD, 2019a). Mongolia and Pakistan are developing countries that are not yet included in international evaluations.

In most of the countries included in our study, countrywide policies aimed at student-centered and inclusive learning have been developed. For instance, in Mongolia, DI and formative assessment have gained a lot of attention through the Mongolia-Cambridge Education Initiative and also, from 2013 on, the “Upright Mongolian child” policies emphasizing equal opportunities and catering to the unique talents of individuals (Government of Mongolia., 2013; Pavlova et al., 2017). In Spain, the government emphasized the need for early diagnosis of problems affecting students’ learning (in the classroom but also regarding access to education) and annual assessment of student performance (Ministerio de Educación y Formación Profesional, 2020). There is also an initiative to provide schools with enough resources for students with specific educational needs. In the Netherlands, knowing how to account for differences between students is part of the standards prospective teachers have to meet before entering the teaching profession, and as such is included in teacher training programs and evaluation criteria for schools (Ministerie van Onderwijs, Cultuur en Wetenschap, 2017). Nevertheless, a lot of Dutch secondary teachers still struggle with fully implementing DI in practice (Van

Casteren et al., 2017). In South Korea, the Master Plan for Educational Welfare with a focus on providing equal opportunities for all students has helped to boost quality of education and to diminish differences in school success caused by students' socio-economic or migrant status (OECD, 2014b). A homeroom teacher functions as a mentor for individual students helping to keep them on track in their development.

Indonesia also has a national policy related to improving teaching quality. However, the country does not have specific policies directed at improving DI or other adaptive teaching practices. Policies directed at improving teaching quality in general have yet to lead to significant improvements (Chang et al., 2014). In Pakistan, there are no specific country-level initiatives aimed at DI either. Studies indicate that Pakistani secondary school teachers rather adopt traditional than students-centred methods of teaching (Andrabi et al., 2013). Whether or not initiatives are employed to boost teaching quality, including DI, teachers in various countries included in our study typically struggle with the implementation of DI (Maulana et al., 2021).

Schools across different countries will probably also vary significantly in the human and material resources they have for accommodating students' learning needs (UNESCO, 2020a). In Indonesia and to a lesser amount in South Korea and Spain, principals reported a shortage of material resources, while shortages in the Netherlands are less pronounced (OECD, 2020). Schools in Mongolia sometimes also experience shortages; for instance, not all schools have access to the internet for pedagogical purposes (UNICEF, 2020). Of the countries in our study, expenditure on education is particularly low in Pakistan and Indonesia (World Bank, 2021). Also, school attendance is a problem in some countries. There are still a lot of children who do not attend secondary education, especially in Pakistan (UNICEF, 2021).

2 Research Questions

In this study, the relationships between personal factors, teaching behaviors and contextual factors and DI are explored across and within different countries. We have different questions guiding this study:

RQ1: Which personal characteristics of the teacher are related to differentiated instruction?

RQ3: Which teaching behavior domains are related to differentiated instruction?

RQ2: To which degree is class size (contextual characteristic) related to differentiated instruction?

RQ4: Are there country-level differences in how characteristics of the teacher, the teaching, and the context are related to differentiated instruction?

Based on the review of the literature, we expect that teaching experience will be positively related to teachers' DI. Since in previous studies other teaching behavior domains were found to be related to DI, we expect to find relations between the other observed teaching behaviors and DI, especially between DI and other relatively

complex teaching behaviors. Class size could be negatively related to teachers' DI, with teachers differentiating more in relatively small classes, although this may not be true for all countries. Since there are large differences in class size across the countries in our sample, the strength of the relation may vary across the different countries. Additionally, there are indications from a Spanish study that females may differentiate more than their male counterparts, but this finding is less clear-cut in the literature. At the school level, some variance may be explained, for instance, because of leadership, practical facilitation of DI and working together with colleagues. At the country level, multiple characteristics may affect how DI is executed and perceived. Policies attempted to stimulate DI like the acts implemented in Mongolia may positively affect DI. In prior studies South Korean teachers were typically found to show high-quality instruction, including DI. In Indonesia and Pakistan, there are no specific country-level initiatives addressing DI, which may lead us to expect less DI in these countries. There may also be between country-differences stemming from differences in how the educational system is set up or how resources are divided. How country-level differences interact with personal- and contextual factors is yet to be explored.

3 Methods

3.1 Sample and Procedure

The current study includes observation data of lessons of a subsample of 1822 teachers in secondary education selected from the data of 4643 teachers from six countries involving Indonesia, the Netherlands, Mongolia, Pakistan, South Korea and Spain. Convenience sampling was used to collect each country sample. All teachers participated on a voluntary basis. Typical lessons of the participating teachers were observed in authentic classroom settings. Data were collected in different years ranging from 2015 to 2020. Observation ratings of one full lesson of each participating teacher were used. More information on the country samples can be found in Maulana et al. (2022).

In the original sample, the number of teachers in both South Korea and the Netherlands was considerably larger than in the other countries (e.g. 2–6 times larger than the sample from Indonesia), which might affect the outcomes. In order to better balance the sample, teachers from these countries were randomly assigned into ten subgroups. We randomly selected a subsample of the subgroups from these two countries for inclusion in the analyses. In the main text, we will present the analyses with the balanced sample of 1822 teachers. The descriptives of the first balanced subsample of in total 1822 teachers included in the main analyses are provided in Table 23.1. The results for two other randomly chosen balanced samples and the unbalanced sample are added to the chapter as supplementary materials (see web version) as a robustness check. More information about the variables can be found in the description of the instruments.

Table 23.1 Descriptives of the balanced sample used in the main analyses per country

	Indonesia	Mongolia	Pakistan	South Korea	Spain	the Netherlands	total
Number of teachers	426	352	373	280	114	277	1822
Number of schools	29	51	20	84	29	163	376
Teacher gender: Female	263	300	179	188	76	169	1175
Teacher subject: alpha*	93	111	162	107	41	113	627
Teacher subject: beta*	184	164	177	124	49	93	791
Teacher subject: gamma*	149	77	34	49	24	71	404
Teacher experience in years(<i>M, sd</i>)	16.2 (9.9)	11.1 (8.6)	6.6 (5.3)	11.4 (8.8)	21.0 (9.5)	3.6 (6.8)	10.9 (9.6)
Teaching behavior: Management (<i>M, sd</i>)	2.8 (0.8)	3.0 (0.5)	2.0 (0.6)	3.2 (0.6)	3.4 (0.6)	3.2 (0.6)	3.2 (0.7)
Teaching behavior: Climate (<i>M, sd</i>)	3.0 (0.6)	3.1 (0.5)	2.0 (0.5)	3.2 (0.6)	3.4 (0.6)	3.3 (0.6)	3.0 (0.7)
Teaching behavior: Instruction (<i>M, sd</i>)	2.7 (0.7)	2.9 (0.5)	1.9 (0.5)	3.1 (0.6)	3.2 (0.6)	3.0 (0.6)	2.9 (0.6)
Teaching behavior: Activation (<i>M, sd</i>)	2.3 (0.6)	2.7 (0.5)	1.9 (0.4)	3.0 (0.6)	3.0 (0.5)	2.5 (0.6)	2.6 (0.6)
Teaching behavior: Learning strategies (<i>M, sd</i>)	2.1 (0.7)	2.7 (0.5)	1.8 (0.5)	2.9 (0.6)	2.8 (0.8)	2.0 (0.7)	2.0 (0.7)
Teaching behavior: Differentiated instruction (<i>M, sd</i>)	1.8 (0.7)	2.3 (0.6)	1.9 (0.5)	2.6 (0.7)	2.2 (0.8)	1.8 (0.7)	2.3 (0.8)
Class size (<i>M, sd</i>)	31.5 (7.8)	26.8 (10.4)	48.0 (14.7)	26.6 (5.8)	17.0 (6.5)	23.1 (5.2)	31.0 (13.3)

* alpha subjects: Native- and foreign language subjects; beta subjects: Mathematics and natural sciences; gamma subjects: Social sciences and humanities

3.2 Instruments

3.2.1 Personal and Contextual Variables

Teachers' gender, school subject and class size were collected by the observers in the classroom. Class size represents the number of students present during the observation. Because of the variety of subjects differing across countries, school subjects were collapsed into three categories: alpha, beta and gamma. Alpha subjects refer to native- and foreign language subjects like Dutch or English. Beta subjects refer to mathematics and natural sciences subjects like science or biology. Gamma subjects refer to social sciences and humanities like history or geography. Subjects in the arts, crafts and physical education were not included in the analyses.

3.2.2 Observation Measure of Teaching Behavior Including Differentiated Instruction

To measure teaching behavior in the six countries, the International Comparative Analysis of Learning and Teaching (ICALT) observation instrument was used (de Grift & Wim, 2014). The instrument consists of 32 high-inferential, observable teaching quality indicators, accompanied by 120 low-inferential observable teaching activities. The differentiation scale of the instrument consists of four high-inferential items like "The teacher offers weaker learners extra study and instruction time" and "The teacher adjusts instruction to relevant inter-learner differences" (see Appendix A for all items and corresponding low inference examples of good practices). Each high-inferential item was rated on a 4-point Likert scale with the following categories: '1 = mostly (predominantly) weak', '2 = more often weak than strong', '3 = more often strong than weak' and '4 = mostly strong'. The sum score of these differentiation items was used as the outcome measure of the study. For all of the countries included in this study the scale reliability is acceptable, ranging from .67 in Pakistan to .84 in South Korea.

The items in the ICALT represent the six domains of teaching behavior discussed in the theoretical section including: safe and stimulating educational climate (4 items), efficient classroom management (4 items), clarity of instruction (7 items), activating teaching (7 items), differentiated instruction (4 items), and teaching learning strategies (6 items). Previous research confirmed the six-factor structure of observed teaching behavior, as well as measurement invariance and applicability of the instrument in secondary schools from different countries (Maulana et al., 2021, 2022). Please refer to Maulana et al. (2021) for examples of items in the other teaching domain-scales.

Trained observers observed a full lesson of each teacher using the ICALT. All observers completed an observer training before they executed the observations. A detailed description of the observer training can be found in Maulana et al. (2021, 2022).

3.3 Analyses

Multilevel regression analyses were used to analyze the relations of different variables with DI in R studio using the packages multilevel (Bliese, 2021; Bliese, 2016), nlmle (Pinheiro et al., 2021), LME4 (Bates et al., 2021; Bates et al., 2015) and sjPlot (Lüdtke, 2021).¹

In order to answer research questions 1–3, we used multilevel modeling by adding personal and contextual variables step-wisely, evaluating the improvement of the model fit as well as the specific influence of different personal and contextual variables. In Model 0, the fixed effect of the school level was added to the model. Then, in Model 1, teachers' gender and experience were added as personal case-mix characteristics of the teacher. After this, teachers' school subject was added to the model (Model 2). In Model 3, indicators of other domains of teaching behavior were added to study the hypothesized relations between teaching behaviors and DI. In Model 4, we added class size as a relevant classroom characteristic. In Model 5, country was added to the equation as a fixed effect. Country was added as a fixed effect instead of as a separate level in the model because there were only 6 countries included in the analyses, which is too limited to treat it as a separate level in the model. Lastly, in order to determine whether the relations between personal and contextual characteristics and DI were affected by the country in which the data was collected, we analyzed Model 4 again splitting the data per country to assess possible country-specific differences.

4 Results

In Fig. 23.1, the results of five different multilevel models are presented. Based on Models 1–4, there is a small, significant effect of gender. The effect of gender is negative for males as compared to females. The estimate becomes insignificant ($p = .056$) in Model 5. There is also a small, positive effect of teaching experience on DI. However, the effect becomes insignificant when the other teaching behavior domains are added into Model 3. The figure further shows that DI is related to classroom management, activating teaching, and teaching learning strategies. Adding the teaching behavior domains improves the model fit most strongly (see Table 23.3). To check whether these results were influenced by the subsample that we used, we compared the findings to results in two other random subsamples and in the unbalanced data (see supplementary materials). Across all random samples, positive relations were found between DI and classroom management, activating teaching and teaching learning strategies. At the country level, significant positive estimates were found for South Korea, Pakistan and (all but one sample) Mongolia. Teaching

¹The analyses were performed in SPSS as well as in R to check comparability. The outcomes were nearly identical (see supplementary materials in the web version of this chapter).

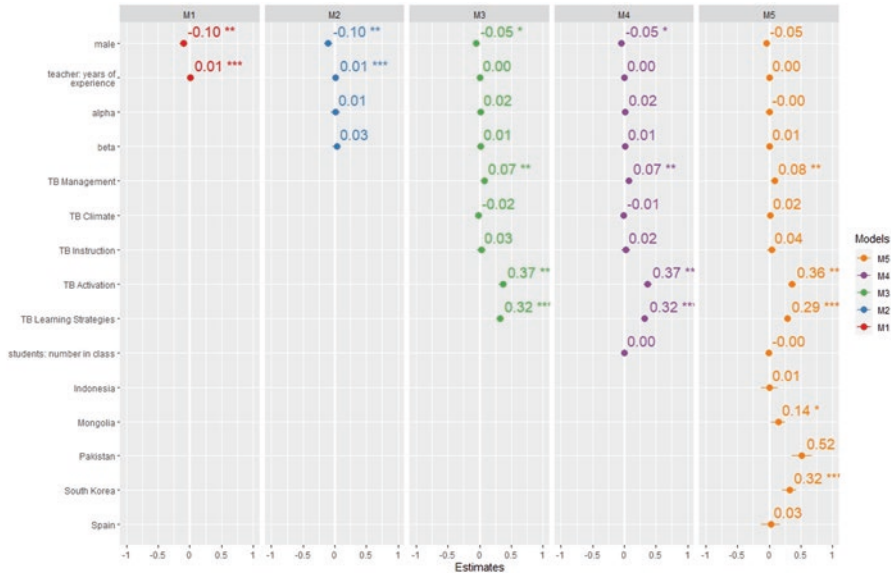


Fig. 23.1 The relations between personal factors, teaching behaviors and contextual factors and DI based on multilevel regression Models 1–5

learning strategies is strongly related to DI in all countries (ranging from $r = .52$ in the Netherlands and Spain to $r = .76$ in Pakistan) as is the quality of activating teaching (ranging from $r = .57$ and $r = .58$ in the Netherlands and Spain respectively to $r = .74$ in South Korea). See Appendix B for the correlations.

The country-level was added to Model 5, showing significantly higher quality DI compared to the Dutch sample for teachers in Pakistan and South Korea, and to a lesser extent Mongolia (see Fig. 23.1 and Table 23.2). The conditional R^2 for Model 5 in Table 23.2 shows that about 70% of the total variance in DI is explained through both fixed and random effects in the model. The ICC indicates that about maximally 33% of this estimated variance could be explained by differences at the school level.

Adding the different countries to the Model significantly improves the model fit (see Table 23.3).

In order to further assess country-level differences regarding how the different personal and contextual characteristics were related to DI, we compared Model 4 across the different countries in Table 23.4.^{2,3} When performing the multilevel analyses for the countries separately, it becomes clear that activating teaching and teaching learning strategies are significant and stable correlates of DI across the different countries. Additionally, in some countries, other teaching behaviors are

²In this case, the full data of South Korea and the Netherlands was used.

³Adding interaction-effects to the full model showed some interactions between variables in the model and different countries, mostly related to the varying effect of experience (see supplementary materials in the web version of the chapter).

Table 23.2 Predictors and estimates of DI based on model 5 of multilevel regression modelling (also presented in the last column of Fig. 23.1).

Predictors	Differentiated instruction		
	Estimates	SE	p-value
Fixed effects			
(intercept)	-0.04	0.09	0.631
Teacher gender male (reference: Female)	-0.05	0.02	0.056
Teacher experience	0.00	0.00	0.840
Teacher subject: alpha (reference: gamma)	-0.00	0.03	0.997
Teacher subject: beta (reference: gamma)	0.01	0.03	0.784
Teaching behavior: Management	0.08	0.03	0.002
Teaching behavior: Climate	0.02	0.03	0.525
Teaching behavior: Instruction	0.04	0.04	0.246
Teaching behavior: Activation	0.36	0.03	<0.001
Teaching behavior: Learning strategies	0.29	0.03	<0.001
Class size	-0.00	0.00	0.530
Country: Indonesia (reference: The Netherlands)	0.01	0.07	0.922
Country: Mongolia (reference: The Netherlands)	0.14	0.06	0.015
Country: Pakistan (reference: The Netherlands)	0.52	0.08	<0.001
Country: South Korea (reference: The Netherlands)	0.32	0.06	<0.001
Country: Spain (reference: The Netherlands)	0.03	0.08	0.744
Random effects			
σ^2 teacher level	0.14		
τ_{00} school level	0.07		
ICC	0.33		
N school	376		
Observations	1822		
Marginal R ² / conditional R ²	0.542 / 0.694		

Table 23.3 Model fit indices of the different multilevel models presented in Fig. 23.1

Model	df	AIC	BIC	logLik	Model Fit	Test	Likelihood Ratio	p-value
Intercept (GLM)	2	3850.1	3861.1	-1923.0	3846.1			
Model 0: Intercept + random effect	3	3071.3	3087.8	-1532.6	3065.3	GLM-0	780.8	<.0001
Model 1: Adding teacher gender + teacher experience	5	3047.8	3075.4	-1518.9	3037.8	0-1	27.4	<.0001
Model 2: Adding teacher subject	7	3050.7	3089.3	-1518.4	3036.7	1-2	1.1	0.5685
Model 3: Adding teaching behavior domains	12	2084.1	2150.2	-1030.0	2060.1	2-3	976.6	<.0001
Model 4: Adding class size	13	2083.9	2155.5	-1029.0	2057.9	3-4	2.1	0.1445
Model 5: Adding countries	18	2031.4	2130.6	-997.7	1995.4	4-5	62.5	<.0001

Table 23.4 Multilevel Model 4 specified for all of the different countries in the sample

<i>Predictors</i>	Indonesia		Mongolia		Pakistan		South Korea		Spain		The Netherlands						
	<i>Estimate</i>	<i>SE</i>	<i>p</i>	<i>Estimate</i>	<i>SE</i>	<i>p</i>	<i>Estimate</i>	<i>SE</i>	<i>p</i>	<i>Estimate</i>	<i>SE</i>	<i>p</i>					
Fixed effects																	
(intercept)	0.19	0.19	0.322	-0.03	0.19	0.887	0.23	0.21	0.14	0.126	0.29	0.45	0.524	0.33	0.09	<0.001	
Teacher gender: Male	-0.01	0.04	0.891	-0.16	0.06	0.006	0.05	0.04	0.202	0.03	-0.09	0.12	0.464	-0.02	0.02	0.380	
Teacher experience	-0.00	0.00	0.666	-0.00	0.00	0.498	0.00	0.00	0.709	0.00	0.003	-0.01	0.01	0.01	0.00	0.003	
Teacher subject: Alpha	-0.01	0.05	0.776	0.02	0.06	0.674	-0.04	0.06	0.466	0.04	0.690	0.14	0.15	0.353	0.03	0.027	
Teacher subject: Beta	-0.03	0.04	0.508	-0.01	0.05	0.885	0.01	0.06	0.889	0.04	0.691	0.19	0.15	0.202	0.03	<0.001	
Teaching behavior: Management	0.02	0.05	0.597	-0.02	0.07	0.753	0.14	0.05	0.004	0.04	0.135	0.13	0.13	0.311	0.03	<0.001	
Teaching behavior: Climate	-0.05	0.06	0.400	0.13	0.06	0.038	0.02	0.04	0.547	0.04	0.093	-0.19	0.12	0.126	0.03	0.220	
Teaching behavior: Instruction	0.02	0.07	0.774	0.20	0.09	0.021	0.11	0.06	0.065	0.05	0.064	-0.14	0.20	0.470	0.03	0.002	
Teaching behavior: Activation	0.30	0.07	<0.001	0.31	0.07	<0.001	0.17	0.06	0.009	0.34	0.05	<0.001	0.81	0.20	<0.001	0.38	<0.001
Teaching behavior: Learning strat.	0.42	0.05	<0.001	0.20	0.07	0.002	0.51	0.05	<0.001	0.26	0.04	<0.001	0.19	0.10	0.080	0.26	<0.001
Class size	0.00	0.00	0.689	-0.00	0.00	0.896	-0.00	0.00	0.016	0.00	0.049	-0.01	0.01	0.554	0.00	0.006	
Random effects																	
σ^2 teacher level	0.14			0.12			0.08		0.13			0.22		0.26			
τ_{00} school level	0.10			0.08			0.00		0.06			0.09		0.03			
ICC	0.41			0.39			0.01		0.3			0.29		0.1			
N schools	29			51			20		142			29		428			
N observations	426			352			373		860			114		2518			
Marginal R2 / conditional R2	0.467 / 0.685			0.418 / 0.646			0.669 / 0.671		0.511 / 0.655			0.423 / 0.589		0.356 / 0.417			

significant predictors like classroom management (the Netherlands and Pakistan), learning climate, and clarity of instruction (Mongolia). In Pakistan, the Netherlands and South Korea, a small negative effect of class size was found indicating that better DI was related to smaller classes. In the Netherlands and South Korea, teaching experience was significantly related to DI. On the other hand, the effect of experience was small and in the reverse direction in the Spanish sample. In the Mongolian sample, a negative effect of gender in favor of females was found. This may be an artefact of the fact that there were few male teachers in the sample. In the Netherlands, alpha and beta subjects were found to be related to higher quality DI as compared to gamma subjects. The percentage of the variance explained at the school level is relatively small, especially in Pakistan and the Netherlands. Overall, there were many commonalities across the countries, but we also found some country-specific influences of personal and contextual factors on DI.

5 Discussion and Conclusions

In this study, we have addressed research questions about how characteristics of the teacher, the teaching and the teaching context are related to teachers' DI and about how these relations differ across countries. Starting with the personal characteristics of the teacher: in our sample, the hypothesis of a small gender effect on DI favoring females was confirmed. Our finding is in line with previous research on gender differences in teaching quality (Fernández-García et al., 2019; Maulana & Helms-Lorenz, 2017). When looking into the country-specific results, the benefit of females is most profound in the Mongolian sample in which only 17% of the teachers was male, which may have affected this finding. Furthermore, a small positive effect of teaching experience was found. This is in line with previous empirical studies (Suprayogi et al., 2017; Van Tassel-Baska et al., 2008; Van der Pers & Helms-Lorenz, 2019) and theoretical assumptions that teachers, overtime, tend to shift their focus from themselves to the learning process of their students (Beijaard et al., 2000; Fuller, 1969). Nevertheless, the positive relation of experience and DI across countries is relatively small and even reversed (experience is negatively related to DI) in Spain. The latter can be caused by the fact that less experienced teachers in Spain tend to be better trained in their initial education and professionalization to address individual students' needs (Fernández-García et al., 2019). The significant relation between experience and DI in Spain and in the Netherlands could also be affected by the fact that the sample in the Netherlands was relatively inexperienced (average experience of 3 years) and the sample in Spain was relatively experienced (average experience of 21 years). Possibly, relations with DI are more profound in these specific groups of teachers. Overall, in our sample, the relations of both gender and experience with DI are small, and they become non-significant when adding teaching behavior indicators to the model. Nevertheless, the fact that they are significant predictors of DI in some of the countries shows that it is interesting to include these personal factors in further investigations. We did not find strong

evidence of differences in relations with DI between school subjects. Only in the Dutch sample, teachers from alpha and beta subjects generally showed higher quality DI than teachers in gamma subjects. More research would be needed to gain insights into differences between specific subjects causing these variations. National-level studies may provide more insight into differences between the execution of DI in specific school subjects within the county.

Indicators of effective teaching behavior were shown to be the strongest correlates of DI in our models. In particular, teachers' ability to manage the classroom, to activate students and to teach about learning strategies were found to be related to teachers' DI. The strong relations between activating teaching, teaching learning strategies and DI are in line with previous studies showing these domains of teaching being clustered together as relatively difficult teaching domains for teachers (Maulana et al., 2021; Maulana et al., 2015; Maulana et al., 2020; Van der Lans et al., 2017). The relatedness of these teaching behaviors can also be traced back to the literature. For instance, expert teachers from the Netherlands stated that they used DI as a means to stimulate students' self-regulative behavior, which is in line with stimulating learning strategies (Keuning et al., 2017; Van Geel et al., 2019). In addition, activating teaching can be connected to DI when teachers deliberately differentiate within the didactical approaches they use to activate students. The relatedness of DI and classroom management was also reported before in literature (Prast et al., 2015). As Tomlinson and Imbeau (2010, preface) write "classroom management is the process of figuring out how to set up and orchestrate a classroom in which students sometimes work as a whole group, as small groups, and as individuals". Teachers who are not able to ensure an orderly and efficient lesson will probably not succeed in flexibly adapting the organization towards DI. But it may also work the other way around; providing students with instruction matching their learning needs may help learners into a state of flow (Csikszentmihályi, 2008) and cultivate a higher sense of competence and autonomy (Deci & Ryan, 1985), which in turn may prevent disorderly behaviors.

Overall, class size was not significantly related to DI over and above the influence of other teaching behavior domains. This is in line with prior findings in secondary education (Brühwiler & Blatchford, 2011). For good teachers who teach in a well-organized, effective manner, some variation in the number of students may not directly affect the quality of their differentiation. Nevertheless, class size was significantly related to DI in some countries. This was the case in Pakistan, South Korea and the Netherlands in which the classes were above average in size; this may make DI more challenging. However, overall, teaching quality seems to be more determining for DI than class size (OECD, 2010).

The variance explained by the school level was limited, even in countries like the Netherlands and South Korea where schools have relatively much autonomy. We did find that teachers in some countries – South Korea, Mongolia and Pakistan – showed higher levels of differentiation relative to teachers in the Netherlands. In Mongolia, classes are relatively heterogeneous and there are specific policy developments aimed at improving individual students' learning processes that may have stimulated teachers' application of DI (Government of Mongolia, 2013; Pavlova et al., 2017). South Korean teachers are typically highly skilled and receive high-quality training and professionalization which may facilitate teaching quality. The

finding that Pakistani teachers showed relatively high-quality DI was somewhat unexpected since educational policies in Pakistan do not specifically address DI and prior research found teachers to show relatively traditional types of teaching (Andrabi et al., 2013). Nevertheless, teachers in Pakistan do have to teach in relatively large, heterogeneous classes with a big spread of learning needs. In such a context, DI seems a logical approach to keep all students on track. Additionally, implementation of DI in Dutch secondary education may be limited since teachers in secondary education may hold the notion that DI is less needed because of the rigorous tracking system (Van Casteren et al., 2017). The fact that teachers in Indonesia showed relatively little high-quality differentiation in reference to other countries is in line with previous studies (Maulana et al., 2021). This may be explained by the fact that DI is not adequately included in educational policies at the country level nor in teacher training or professionalization programs. The fact that Spanish teachers did not show higher quality DI than teachers in the Netherlands may partly be affected by the relatively experienced sample in this study. In Spain, inexperienced teachers were found to implement DI better than more experienced counterparts (Fernández-García et al., 2019). Also, policies regarding attending to individual differences are relatively new and it may take some time before they affect daily classroom practices.

Although we can hypothesize about country-specific circumstances that may explain differences in correlates of DI, more in-depth studies are needed to verify such influences. One finding that is consistent throughout our study though, is that across and within the participating countries, teaching quality in other domains of teaching – particularly activating instruction and learning strategies – is related to the implementation of DI.

Scientific and Practical Implications On the scientific level, the fact that activating teaching and teaching learning strategies are positively related to DI is in line with a stage-like framework of teaching in which these relatively difficult domains of teaching cluster together (Maulana et al., 2021; Van der Lans et al., 2017). The relatedness across the domains could also adhere to the idea that these teaching domains can be clustered into a broader overarching domain aimed at student-centered teaching or student-support (compare the model of Praetorius et al., 2018).

On a practical note, the relatedness between different domains of teaching may imply that educators aiming to stimulate DI are best off targeting a broad development of teaching behaviors that may facilitate DI. For example, (prospective) teachers could be taught how to manage the classroom well in order to teach them skills useful for managing different instructional routes. Alternatively, related teaching behaviors may be taught in interaction. For instance, teacher educators could prompt teachers to activate their students by using differentiated activating approaches suitable to students' learning needs. By helping their students to monitor their own learning and by encouraging the use of learning strategies differentiated to students' needs, teachers could connect the dots between differentiation and self-regulated learning. Lastly, we found that personal and contextual factors could affect the implementation of DI to a certain extent. Teaching does not happen in a vacuum and professionalization initiatives should thus take the teachers' characteristics and context into account.

Limitations Although this chapter explores the characteristics of the teacher, the teaching process and the context of the classroom, school, and country with observations from a broad range of educational contexts, there are some limitations. First, although observation measures are suitable to capture a lot of information in authentic situations, the observation instrument used in this study does not capture all aspects of DI. The concept was measured using certain specific indicators focusing on convergent differentiation (aimed at supporting weaker students) and on differentiation of instructions and processing. Other forms of DI such as differentiation of learning materials, differentiating the end product and making adaptations in the learning environment are underrepresented. Future refinement of the instrument could help to capture a more comprehensive operationalization of DI. In addition, the observational data do not give insights in the reasoning of the teachers when implementing DI. Further research is needed to get more insight in the why's and how's of the teaching behavior (Gheysens et al., 2021; Vantieghem et al., 2020). Additionally, although the lesson observations give valuable insights into classrooms across the globe, only one lesson of each teacher was included. Across the sample, the mean scores presumably give a good indication of the average DI of teachers. Nevertheless, data from one lesson may be less suitable for reflection on individual qualities of teachers. In studies that aim to give insights on the individual level, more lesson observations should be included (Van der Lans et al., 2016).

Secondly, although the data from the individual countries are sufficiently large and relatively representative, teachers participated on a voluntary basis. This means that the current sample may not include specific groups of teachers needed for making inferences at the country level. Hence, caution against the generalization of findings to the country level is warranted until replication studies with broader and more representative samples are available.

Lastly, only a limited number of variables about personal- and contextual factors were collected because of practical reasons. There are relevant variables that were not included into our study like heterogeneity of the class (Tomlinson et al., 2003), team collaboration in the school (Smit & Humpert, 2012), lesson materials and curriculum (Van Geel et al., 2019), teacher beliefs and self-efficacy for implementing DI (Suprayogi et al., 2017; Whitley et al., 2019) and professional vision (Gheysens et al., 2021; Vantieghem et al., 2020). This study offers an insightful starting point, but further studies including more personal-, pedagogical-didactical and contextual characteristics are needed to shed more light on how teachers' DI is related to personal characteristics, teaching and context.

Funding This work was supported by the Dutch scientific funding agency (NRO) under Grant number 405-15-732; the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea under Grant number NRF-2017S1A5A2A03067650; and the Directorate General of Higher Education of Indonesia under Grant number 04/SP2H/DRPM/LPPM-UNJ/III/2019; Institute of Educational Research and Innovation of the University of Oviedo (INIE), under grant INIE-19- MOD C-1.

Appendix A

The DI-scale of the ICALT observation instrument

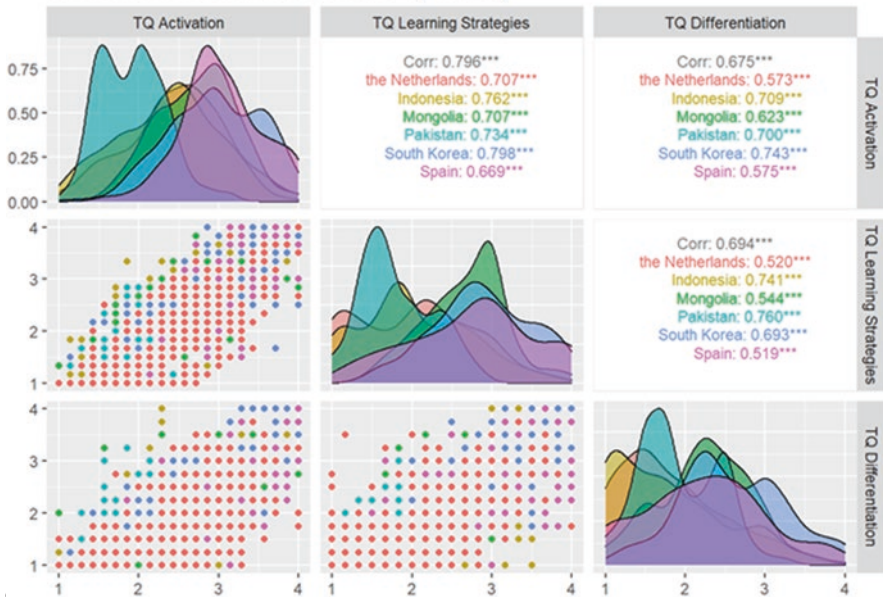
.		Indicator: The teacher...	Examples of good practice: <i>The teacher ...</i>	Observed	
Differentiated instruction	23	...evaluates whether the lesson aims have been reached	1 2 3 4 ...evaluates whether the lesson aims have been reached	0 1	
			...evaluates learners' performance	0 1	
	24	...offers weaker learners extra study and instruction time	1 2 3 4	...gives weaker learners extra study time	0 1
				...gives weaker learners extra instruction time	0 1
				...gives weaker learners extra exercises/practices	0 1
				...gives weaker learners 'pre- or post-instruction'	0 1
	25	...adjusts instructions to relevant inter-learner differences	1 2 3 4	...puts learners who need little instructions (already) to work	0 1
				...gives additional instructions to small groups or individual learners	0 1
				...does not simply focus on the average learner	0 1
	26	...adjusts the processing of subject matter to relevant inter-learner differences	1 2 3 4	...distinguishes between learners in terms of the length and size of assignments	0 1
				...allows for flexibility in the time learners get to complete assignments	0 1
				...lets some learners use additional aids and means	0 1

Note. *The ICALT instrument is freely available upon request. However, do note that use of the instrument requires extensive and proper training. Examples of high and low inference indicators of the other teaching behavior domains can be found in the paper of Maulana et al. (2021)*

Appendix B

Correlations between DI and the 'activating teaching'- and 'teaching learning strategies' scale of the ICALT across the countries in our

correlation 3 most difficult domains by country



References

- Andrabi, T., Das, J., & Khwaja, A. I. (2013). Students today, teachers tomorrow: Identifying constraints on the provision of education. *Journal of Public Economics*, 100(1), 1–14. <https://doi.org/10.1016/j.jpubeco.2012.12.003>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). *Fitting linear mixed-effects models using lme*, 4.67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2021). *Lme4: Linear mixed-effects models using 'eigen' and S4* (R package version 1.1-27 ed.) <https://cran.r-project.org/web/packages/lme4>
- Beijaard, D., Verloop, N., & Vermunt, J. D. (2000). Teachers' perceptions of professional identity: An exploratory study from a personal knowledge perspective. *Teaching and Teacher Education*, 16(7), 749–764. [https://doi.org/10.1016/S0742-051X\(00\)00023-8](https://doi.org/10.1016/S0742-051X(00)00023-8)
- Berliner, D. C. (2004). Describing the behavior and documenting the accomplishments of expert teachers. *Bulletin of Science, Technology, & Society*, 24(3), 200–212. <https://doi.org/10.1177/0270467604265535>
- Betts, J. R., & Shkolnik, J. T. (1999). The behavioral effects of variations in class size: The case of math teachers. *Educational Evaluation and Policy Analysis*, 21(2), 193–213. <https://doi.org/10.2307/1164300>
- Blatchford, P., Bassett, P., & Brown, P. (2011). Examining the effect of class size on classroom engagement and teacher-pupil interaction: Differences in relation to pupil prior attainment and primary vs. secondary schools. *Learning and Instruction*, 21(6), 715–730. <https://doi.org/10.1016/j.learninstruc.2011.04.001>
- Bliese, P. (2016). *Multilevel modeling in R (2.6): A brief introduction to R, the multilevel package and the nlme package*. <https://CRAN.R-project.org/package=multilevel>
- Bliese, P. (2021). *Multilevel: Multilevel functions*. <https://CRAN.R-project.org/package=multilevel>

- Bondie, R. S., Dahnke, C., & Zusho, A. (2019). How does changing “one-size-fits-all” to differentiated instruction affect teaching? *Review of Research in Education*, 43, 336–362. <https://doi.org/10.3102/0091732X18821130>
- Brühwiler, C., & Blatchford, P. (2011). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning and Instruction*, 21(1), 95–108. <https://doi.org/10.1016/j.learninstruc.2009.11.004>
- Chang, M. C., et al. (2014). *Teacher reform in Indonesia: The role of politics and evidence in policy making. Directions in development—Human development*. World Bank.
- Chun, S., Ulziisaikhan, G., Kim, D., Amarjargal, A., & Kim, J. (2020). A study on the applicability of ICALT for comparing teaching skills between secondary schools in Korea and Mongolia. *The SNU Journal of Education Research*, 29(1), 53–75.
- Coubergs, C., Struyven, K., Vanthournout, G., & Engels, N. (2017). Measuring teachers' perceptions about differentiated instruction: The DI-quest instrument and model. *Studies in Educational Evaluation*, 53, 41–54. <https://doi.org/10.1016/j.stueduc.2017.02.004>
- Csikszentmihályi, M. (2008). *Flow: The psychology of optimal experience*. HarperCollins Publishers.
- de Grift, V., & Wim, J. C. M. (2014). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement*, 25(3), 295–311. <https://doi.org/10.1080/09243453.2013.794845>
- De Jager, T. (2013). Guidelines to assist the implementation of differentiated learning activities in south African secondary schools. *International Journal of Inclusive Education*, 17(1), 80–94. <https://doi.org/10.1080/13603116.2011.580465>
- De Jager, T. (2017). Perspectives of teachers on differentiated teaching in multi-cultural south African secondary schools. *Studies in Educational Evaluation*, 53, 115–121. <https://doi.org/10.1016/j.stueduc.2016.08.004>
- de Jager, T., Coetzee, M. J., Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2017). Profile of south African secondary-school teachers' teaching quality: Evaluation of teaching practices using an observation instrument. *Educational Studies*, 43(4), 410–429. <https://doi.org/10.1080/03055698.2017.1292457>
- De Neve, D., & Devos, G. (2016). The role of environmental factors in beginning teachers' professional learning related to differentiated instruction. *School Effectiveness and School Improvement*, 27(4), 557–579. <https://doi.org/10.1080/09243453.2015.1122637>
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum.
- Deunk, M. I., Smale-Jacobse, A. E., de Boer, H., Doolaard, S., & Bosker, R. J. (2018). Effective differentiation practices: A systematic review and meta-analysis of studies on the cognitive effects of differentiation practices in primary education. *Educational Research Review*, 24, 31–54. <https://doi.org/10.1016/j.edurev.2018.02.002>
- Education policy and data center. (2018). *National Education Profile 2018 update*. Education policy and data center. https://www.epdc.org/sites/default/files/documents/EPDC_NEP_2018_Mongolia.pdf
- Fernández-García, C. M., Maulana, R., Inda-Caro, M., Helms-Lorenz, M., & García-Pérez, O. (2019). Student perceptions of secondary education teaching effectiveness: General profile, the role of personal factors, and educational level. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.00533>
- Fuller, F. F. (1969). Concerns of teachers: A developmental conceptualization. *American Educational Research Journal*, 6(2), 207–226.
- Gamoran, A. (1992). The variable effects of high school tracking. *American Sociological Review*, 57(6), 812–828. <https://doi.org/10.2307/2096125>
- Gay, G. (2013). Teaching to and through cultural diversity. *Curriculum Inquiry*, 43(1), 48–70. <https://doi.org/10.1111/curi.12002>
- Gheysens, E., Consuegra, E., Engels, N., & Struyven, K. (2021). Creating inclusive classrooms in primary and secondary schools: From noticing to differentiated practices. *Teaching and Teacher Education*, 100, 103210. <https://doi.org/10.1016/j.tate.2020.103210>

- Goddard, Y. L., Neumerski, C. M., Goddard, R. D., Salloum, S. J., & Berebitsky, D. (2010). A multilevel exploratory study of the relationship between teachers' perceptions of principals' instructional support and group norms for instruction in elementary schools. *Elementary School Journal*, 111(2), 336–357. <https://doi.org/10.1086/656303>
- Government of Mongolia. (2013). *Upright Mongolian child program 2013–2016*. Government Protocol No.295.
- Grossman, P. L., & Stodolsky, S. S. (1995). Content as context: The role of school subjects in secondary school teaching. *Educational Researcher*, 24(8), 5–23. <https://doi.org/10.2307/1176887>
- Hayden, H. E., Rundell, T. D., & Smyntek-Gworek, S. (2013). Adaptive expertise: A view from the top and the ascent. *Teaching Education*, 24(4), 395–414. <https://doi.org/10.1080/10476210.2012.724054>
- Helms-Lorenz, M., & Visscher, A. J. (2021). Unravelling the challenges of the data-based approach to teacher improvement. *School Effectiveness and School Improvement.*, 33, 125–147. <https://doi.org/10.1080/09243453.2021.1946568>
- Hendayana, S., Supriatna, A., & Imansyah, H. (2010). Indonesia's issues and challenges on quality improvement of mathematics and science education. *Journal of International Cooperation in Education*, 4(2), 41–51.
- Hertberg-Davis, H., & Brighton, C. M. (2006). Support and sabotage: Principals' influence on middle school teachers' responses to differentiation. *Journal of Secondary Gifted Education*, 17(2), 90–102. <https://doi.org/10.4219/jsge-2006-685>
- Keuning, T., van Geel, M., Frèrejean, J., van Merriënboer, J., Dolmans, D., & Visscher, A. J. (2017). Differentiëren bij rekenen: Een cognitieve taakanalyse van het denken en handelen van basisschoolleerkrachten [differentiated instruction for mathematics: A cognitive task analysis into primary school teachers' reasoning and acting]. *Pedagogische Studiën*, 94(3), 160–181.
- Kyriakides, L., Creemers, B. P. M., & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education*, 25(1), 12–23. <https://doi.org/10.1016/j.tate.2008.06.001>
- Lavania, M., & Nor, F. B. M. (2020). Barriers in differentiated instruction. A systematic review of the literature. *Journal of Critical Reviews (discontinued)*, 7(6), 293–297.
- Loreman, T., Forlin, C., Chambers, D., Sharma, U., & Deppeler, J. (2014). Conceptualising and measuring inclusive education. In C. Forlin & T. Loreman (Eds.), *Measuring inclusive education. (International perspectives on inclusive education volume 3)* (pp. 3–17).
- Lüdtke, D. (2021). *sjPlot: Data visualization for statistics in social science* (R package version 2.8.7 ed.). <https://cran.r-project.org/web/packages/sjPlot>
- Maulana, R., Helms-Lorenz, M., & Van, de Grift (2017). Validating a model of effective teaching behaviour of pre-service teachers. *Teachers and Teaching*, 23(4), 471–493. doi: <https://doi.org/10.1080/13540602.2016.1211102>
- Maulana, R., Helms-Lorenz, M., & Van de Grift, W. (2015). Development and evaluation of a questionnaire measuring pre-service teachers' teaching behaviour: A rasch modelling approach. *School Effectiveness and School Improvement*, 26(2), 169–194. <https://doi.org/10.1080/09243453.2014.939198>
- Maulana, R., Smale-Jacobse, A. E., Helms-Lorenz, M., Chun, S., & Lee, O. (2020). Measuring differentiated instruction in The Netherlands and South Korea: Factor structure equivalence, correlates, and complexity level. *European Journal of Psychology of Education*, 35, 881–909. <https://doi.org/10.1007/s10212-019-00446-4>
- Maulana, R., André, S., Helms-Lorenz, M., Ko, J., Chun, S., Shahzad, A., et al. (2021). Observed teaching behaviour in secondary education across six countries: Measurement invariance and indication of cross-national variations. *School Effectiveness and School Improvement*, 32(1), 64–95. <https://doi.org/10.1080/09243453.2020.1777170>
- Maulana, R., Helms-Lorenz, M., Moorer, P., Smale-Jacobse, A., & Feng, X. (2022). *Differentiated instruction in teaching from the international perspective: Cross-country comparisons, development over time, and the importance for students' Academic Engagement*. University of Groningen.

- Ministerie van Onderwijs, Cultuur en Wetenschap. (2017, March). Besluit van 16 maart 2017 tot wijziging van het Besluit bekwaamheidseisen onderwijspersoneel en het Besluit bekwaamheidseisen onderwijspersoneel BES in verband met de herijking van de bekwaamheidseisen voor leraren en docenten, Staatsblad, 148.
- Ministerio de Educación y Formación Profesional. (2020). Ley Orgánica 3/2020, de 29 de diciembre, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación (LOMLOE). *Boletín Oficial del Estado*, 340, 122868–122952.
- Ministry of Education and Science. (2021, January 20). <https://www.meds.gov.mn/post/64833>
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016: International results in reading. International Association for the Evaluation of Educational Achievement*. [http://search.ebscohost.com.proxy-ub.rug.nl/login.aspx?direct=true & db=eric & AN=ED580353 & site=ehost-live & scope=site](http://search.ebscohost.com.proxy-ub.rug.nl/login.aspx?direct=true&db=eric&AN=ED580353&site=ehost-live&scope=site)
- Mullis, I. V., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019. International results in mathematics and science*. Lynch School of Education and Human Development.
- OECD. (2010). *PISA 2009 results (volume IV): What makes a school successful?* OECD Publishing.
- OECD. (2011). *PISA in focus. School autonomy and accountability: Are they related to student performance?* doi:<https://doi.org/10.1787/5k9h362kcx9w-en>.
- OECD. (2012). *Equity and quality in education. Supporting disadvantaged students and schools*. OECD Publishing.
- OECD. (2014a). How many students are in each classroom?. In *Education at a glance 2014: Highlights*, OECD Publishing. doi:https://doi.org/10.1787/eag_highlights-2014-24-en.
- OECD. (2014b). *Lessons from PISA for Korea, strong performers and successful reformers in education*. OECD Publishing.
- OECD. (2016a). *Education policy outlook Korea*. OECD Publishing.
- OECD. (2016b). *Netherlands 2016. Foundations for the Future*. OECD Publishing. <https://doi.org/10.1787/9789264257658-en>
- OECD. (2016c). *PISA 2015 results (volume I): Excellence and equity in education*. OECD Publishing. <https://doi.org/10.1787/9789264266490-en>
- OECD. (2019a). *PISA 2018 results (volume I). What students know and can do*. OECD Publishing.
- OECD. (2019b). *TALIS 2018 results (volume I). Teachers and school leaders as lifelong learners*. OECD Publishing.
- OECD. (2020). *PISA 2018 results (volume V): Effective policies, successful schools*. OECD Publishing. <https://doi.org/10.1787/ca768d40-en>
- OECD. (2021). *Education at a glance 2021: OECD indicators*. OECD Publishing. <https://doi.org/10.1787/b35a14e5-en>
- Opendakker, M., & Van Damme, J. (2007). Do school context, student composition and school leadership affect school practice and outcomes in secondary education? *British Educational Research Journal*, 33(2), 179–206. <https://doi.org/10.1080/01411920701208233>
- Opendakker, M., Maulana, R., & den Brok, P. (2012). Teacher–student interpersonal relationships and academic motivation within one school year: Developmental changes and linkage. *School Effectiveness and School Improvement*, 23(1), 95–119. <https://doi.org/10.1080/09243453.2011.619198>
- Pavlova, M., Lee, J. C., & Maclean, R. (2017). *Transitions to post-school life: Responsiveness to individual, social and economic needs*. Springer.
- Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2021). *Nlme: Linear and nonlinear mixed effects models* (R package version 3.1-152 ed.) <https://CRAN.R-project.org/package=nlme>
- Pozas, M., Letzel, V., & Schneider, C. (2020). Teachers and differentiated instruction: Exploring differentiation practices to address student diversity. *Journal of Research in Special Educational Needs*, 20(3), 217–230. <https://doi.org/10.1111/1471-3802.12481>
- Praetorius, A., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of three basic dimensions. *ZDM*, 50(3), 407–426. <https://doi.org/10.1007/s11858-018-0918-4>
- Prast, E. J., Van de Weijer-Bergsma, E., Kroesbergen, E. H., & Van Luit, J. E. H. (2015). Readiness-based differentiation in primary school mathematics: Expert recommendations and teacher self-assessment. *Frontline Learning Research*, 3(2), 90–116. <https://doi.org/10.14786/flr.v3i2.163>

- Reynolds, D., Creemers, B. P. M., Stringfield, S., Teddlie, C., & Schaffer, E. (2002). *World class schools: International perspectives in school effectiveness*. Routledge Falmer.
- Rock, M. L., Gregg, M., Ellis, E., & Gable, R. A. (2008). REACH: A framework for differentiating classroom instruction. *Preventing School Failure, 52*(2), 31–47. <https://doi.org/10.3200/PSFL.52.2.31-47>
- Roiha, A. (2014). Teachers' views on differentiation in content and language integrated learning (CLIL): Perceptions, practices and challenges. *Language and Education, 28*(1), 1–18. <https://doi.org/10.1080/09500782.2012.748061>
- Schipper, T., Goei, S. L., de Vries, S., & van Veen, K. (2017). Professional growth in adaptive teaching competence as a result of lesson study. *Teaching and Teacher Education, 68*, 289–303. <https://doi.org/10.1016/j.tate.2017.09.015>
- Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research Design in Disentangling Meta-Analysis Results. *Review of Educational Research, 77*(4), 454–499. <https://doi.org/10.3102/0034654307310317>
- Smale-Jacobse, A. E., Meijer, A., Helms-Lorenz, M., & Maulana, R. (2019). Within-class differentiation in secondary education: A systematic review of research evidence. *Frontiers in Psychology, 10*(2366). <https://doi.org/10.3389/fpsyg.2019.02366>
- Smit, R., & Humpert, W. (2012). Differentiated instruction in small schools. *Teaching and Teacher Education, 28*(8), 1152–1162. <https://doi.org/10.1016/j.tate.2012.07.003>
- Steenbergen-Hu, S., Makel, M. C., & Olszewski-Kubilius, P. (2016). What one hundred years of research says about the effects of ability grouping and acceleration on K–12 students' academic achievement: Findings of two second-order meta-analyses. *Review of Educational Research, 86*(4), 849–899. <https://doi.org/10.3102/0034654316675417>
- Sun, Y., & Xiao, L. (2021). Research trends and hotspots of differentiated instruction over the past two decades (2000-2020): A bibliometric analysis. *Educational Studies.*, 1–17. <https://doi.org/10.1080/03055698.2021.1937945>
- Suprayogi, M. N., Valcke, M., & Godwin, R. (2017). Teachers and their implementation of differentiated instruction in the classroom. *Teaching, & Teacher Education, 67*, 291–301. <https://doi.org/10.1016/j.tate.2017.06.020>
- Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners* (2nd ed.). ASCD.
- Tomlinson, C. A., & Imbeau, M. B. (2010). *Leading and managing a differentiated classroom*. ASCD.
- Tomlinson, C. A., Brighton, C., Hertzberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., et al. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *Journal for the Education of the Gifted, 27*(2–3), 119–145. <https://doi.org/10.1177/016235320302700203>
- UNESCO. (2017). *A guide for ensuring inclusion and equity in education*. United Nations Educational, Scientific and Cultural Organization.
- UNESCO. (2019). *Education in Mongolia: A country report*. United Nations Educational, Scientific and Cultural Organization.
- UNESCO. (2020a). *Inclusion and education: All means all*. United Nations Educational, Scientific and Cultural Organization.
- UNESCO. (2020b). *Mongolia education policy review: Towards a lifelong learning system*. United Nations Educational, Scientific and Cultural Organization.
- UNESCO. (2021). <http://data.uis.unesco.org/>
- UNICEF. (2020). *Review of educational management systems (EMIS) that track individual student data*. UNICEF.
- UNICEF. (2021, April). <https://data.unicef.org/topic/education/secondary-education/#:~:text=Despite%20tremendous%20progress%2C%20startling%20differences,the%20richest%20and%20the%20poorest>
- United States Agency for International Development (2006). *Situation analysis of teacher education: Towards a strategic framework for teacher education and professional development*. .
- Van Bergen, K., Paulussen-Hoogbeem, M., Mack, A., Rossing, H., & Van der Ploeg, Y. (2016). *Groeps-grootte in het VO*. Regioplan.

- Van Casteren, W., Bendig-Jacobs, J., Wartenbergh-Cras, F., Van Essen, M., & Kurver, B. (2017). *Differentiëren en Differentiatievaardigheden in Het Voortgezet Onderwijs*. ResearchNed.
- Van der Lans, R. M., Van de Grift, W., Van Veen, K., & Fokkens-Bruinsma, M. (2016). Once is not enough: Establishing reliability criteria for feedback and evaluation decisions based on classroom observations. *Studies in Educational Evaluation*, 50, 88–95. <https://doi.org/10.1016/j.stueduc.2016.08.001>
- Van der Lans, R. M., Van de Grift, W., & Van Veen, K. (2017). Individual differences in teacher development: An exploration of the applicability of a stage model to assess individual teachers. *Learning and Individual Differences*, 58, 46–55. <https://doi.org/10.1016/j.lindif.2017.07.007>
- Van der Lans, R., Cremers, J., Klugkist, I., & Zwart, R. (2020). Teachers' interpersonal relationships and instructional expertise: How are they related? *Studies in Educational Evaluation*, 66, 100902. <https://doi.org/10.1016/j.stueduc.2020.100902>
- Van der Pers, M., & Helms-Lorenz, M. (2019). Regional school context and teacher characteristics explaining differences in effective teaching behaviour of beginning teachers in The Netherlands. *School Effectiveness and School Improvement*, 30(2), 1–24. <https://doi.org/10.1080/09243453.2019.1592203>
- Van Geel, M., Keuning, T., Frèrejean, J., Dolmans, D., van Merriënboer, J., & Visscher, A. J. (2019). Capturing the complexity of differentiated instruction. *School Effectiveness, & School Improvement*, 30(1), 51–67. <https://doi.org/10.1080/09243453.2018.1539013>
- Van Petegem, K., Creemers, B. P. M., Rossel, Y., & Aelterman, A. (2005). Relationships between teacher characteristics, interpersonal teacher behavior and teacher wellbeing. *Journal of Classroom Interaction*, 40(2), 34–43. <https://www.jstor.org/stable/23870662>
- Van Tassel-Baska, J., Feng, A., MacFarlane, B., Heng, M. A., Teo, C. T., Wong, M. L., et al. (2008). A cross-cultural study of teachers' instructional practices in Singapore and the United States. *Journal for the Education of the Gifted*, 31(3), 214–239. <https://doi.org/10.4219/jeg-2008-770>
- Vantieghem, W., Roose, I., Gheysens, E., Griful-Freixenet, J., Keppens, K., Vanderlinde, R., et al. (2020). Professional vision of inclusive classrooms: A validation of teachers' reasoning on differentiated instruction and teacher-student interactions. *Studies in Educational Evaluation*, 67, 100912. <https://doi.org/10.1016/j.stueduc.2020.100912>
- Vygotskii, L. S., & Cole, M. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wan, S. W. (2014). Differentiated instruction: Are Hong Kong in-service teachers ready? *Teachers and Teaching*, 23(3), 284–311. <https://doi.org/10.1080/13540602.2016.1204289>
- Wang, M. T., & Degol, J. L. (2016). School climate: A review of the construct, measurement, and impact on student outcomes. *Educational Psychology Review*, 28, 315–352. <https://doi.org/10.1007/s10648-015-9319-1>
- Whitley, J., Gooderham, S., Duquette, C., Orders, S., & Bradley Cousins, J. (2019). Implementing differentiated instruction: A mixed-methods exploration of teacher beliefs and practices. *Teachers and Teaching*, 25(8), 1043–1061. <https://doi.org/10.1080/13540602.2019.1699782>
- World Bank. (2015). Indonesia teacher certification and beyond: An empirical evaluation of the teacher certification program and education quality improvements in Indonesia. <https://openknowledge.worldbank.org/handle/10986/24433>
- World bank. (2021, September). *Government expenditure on education, total (% of GDP)*. <https://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS>

Annemieke Smale-Jacobse Annemieke Smale-Jacobse authored this study while working as an assistant professor at the Department of Teacher Education of the University of Groningen. She was involved in different courses on teaching methodology and pedagogy and research on different aspects of teaching and learning. Her research interests include differentiated instruction and adaptive teaching, metacognition, problem solving, reading comprehension and professionalization of teachers. Currently, she is working as an advisor for education and research at the Hanze University of Applied Sciences. email: a.e.smale@pl.hanze.nl

Peter Moorer drs. Peter Moorer is a former researcher from the Department of Teacher Education of the University of Groningen who has particularly been involved in data management and analyses together with other members of the research team. His research interests are in theoretical sociology, psychology and economics. To solve complex theoretical issues, he has specialized in advanced statistical analyses (GLM, GLMM, SEM and Data Mining).

Ridwan Maulana is an associate professor at the Department of Teacher Education, University of Groningen, the Netherlands. His major research interests include teaching and teacher education, factors influencing effective teaching, methods associated with the measurement of teaching, longitudinal research, cross-country comparisons, effects of teaching behaviour on students' motivation and engagement, and teacher professional development. He has been involved in various teacher professional development projects including the Dutch induction programme and school–university-based partnership. He is currently a project leader of an international project on teaching quality involving countries from Europe, Asia, Africa, Australia, and America. He is a European Editor of Learning Environments Research journal, a SIG leader of Learning Environments of American Educational Research Association, and chair of the Ethics Commission of the Teacher Education.

Michelle Helms-Lorenz is an Associate Professor at the Department of Teacher Education, University of Groningen, The Netherlands. Her research interest covers the cultural specificity versus universality (of behaviour and psychological processes). This interest was fed by the cultural diversity in South Africa, where she was born and raised. Michelle's second passion is education, the bumpy road toward development. Her research interests include teaching skills and well-being of beginning and pre-service teachers and effective interventions to promote their professional growth and retention.

Carmen-María Fernández-García, PhD, Associate Professor at the Department of Educational Sciences at the University of Oviedo (Spain). She has received research grants from the Spanish Ministry of Education. She is member of the Spanish Society of Comparative Education, the Spanish Society of Pedagogy and the ASOCED Research Group. Her major research interests involve teaching and teacher education, learning and instruction, gender and comparative education. She has published several academic papers on these topics. Currently she is joining an international project investigating teaching behavior and student outcomes across countries, the ICALT3 Project coordinated by the University of Groningen. email: fernandezcarmen@uniovi.es

Mercedes Inda-Caro, PhD, Associate Professor at the University of Oviedo (Spain). She previously worked as a training support counselor in a public school as part of her FICYT scholarship training (1997) and as Child Educator for the Principality of Asturias within the Ministry of Social Services in two periods (1996/2000). Her PhD dealt with the concept of personality disorders. Currently, she is working on three lines of research: family and gender, teacher and teaching-learning education, and gender and technology studies, as a member of the ASOCED Research Group. She has several publications in scientific journals. email: indamaria@uniovi.es

Prof. Seyoung Chun is Professor Emeritus of Education at Chungnam National University, one of the major national universities in Daejeon, Korea. He received his education and Ph.D. from Seoul National University, South Korea, and has been actively engaged in education policy research and has held several key positions such as Secretary of Education to the President and CEO of KERIS. He founded the Smart Education Society in 2013, and has led many projects and initiatives for the paradigm shift of education in the digital era. Since his early career at the Korean National Commission for UNESCO, he has participated in many international cooperation projects and worked for several developing countries such as Nicaragua, Honduras, Cambodia, etc. *Education Miracle in the Republic of Korea* is the latest book to be published as a summary of his academic life.

Dr. Abid Shahzad is the Founding Director of the International Linkages at the Islamia University of Bahawalpur (IUB). He earned his PhD degree in Educational Sciences from Ghent University Belgium. Currently, he is serving as an Associate Professor and Chair of the Department of Educational Leadership and Management. He has presented his research papers and participated in international workshops and seminars in more than twenty countries. He is also the founder of the International Conference on Teaching and Learning (ICOTAL) and International STEMS conference that are held annually at the Islamia University of Bahawalpur. He has engaged a number of renowned international universities, research institutes and educationists in the ICOTAL and STEMS conferences. He is regularly organizing international training workshops for research students at the Faculty of Education. He is actively signing MoUs with international academic and research partners. He is a national trainer of faculty capacity building program in Pakistan.

Okhwa Lee is a professor (emeritus) at the Department of Education, Chungbuk National University, South Korea, and CEO of SmartSchool (Ltd). Okhwa Lee is a specialist in educational technology and a practitioner in pre-service teacher education. She is a pioneer of software education, e-learning, and smart education in South Korea. She has been a member of the Presidential Educational Reform Committee and the Presidential e-Government of South Korea including various department committees. She has collaborated in the global society such as European Erasmus mobility programme, UNESCO, OECD and in the Korean government ODA (Official Development Assistant) programmes for Nicaragua, Cambodia, Myanmar, Nigeria, Vietnam, Thailand, Philippines and Ethiopia.

Amarjargal Adiyasuren is a lecturer at Mongolian National University of Education. She formerly worked in Teachers' Professional Development Institute and Curriculum Reform Unit affiliated to Ministry of Education and Science. She worked in various national research projects related to school management, curriculum, pedagogy and assessment. She has been involved in comparative study of assessment of transversal skills with the Network on Education Quality Monitoring in the Asia-Pacific in the UNESCO Asia-Pacific and the Brookings Institution of the USA. She holds bachelor and master degree in Education from the University of Tokyo. email: a.amarjargal@gmail.com

Yulia Iridayanti obtained her first degree in Biology Education and PhD in Biology. She is currently a Senior Lecturer and researcher at the Biology and Biology Education Department, Universitas Negeri Jakarta [State University of Jakarta], Indonesia. Since 2001, she has been working together with the Teacher Education Department of University of Groningen, the Netherlands, on the project about teaching quality and student academic motivation from the international perspective (ICALT3/Differentiation project, Principal investigator Indonesia). She is interested in helping teachers to improve their teaching quality and student differences in their learning needs, motivation, and learning style.

Ulziisaikhan Galindev is a senior lecturer in The Department of Educational Administration, Mongolian National University of Education. He received his master and doctoral degrees in Educational administration from Chungnam National University, South Korea. His current research interests and expertise cover education finance, education policy and teacher professional development. Email: ulziisaikhan@msue.edu.mn

Nurul Fadhilah works part-time as a lecturer at the Department of Biostatistic and Population, University of Indonesia. She has been actively involved in the international project called ICALT3/ Differentiation as an expert observer and as co-investigator for Indonesia. Currently, she is engaged in research projects related to digital health within the health informatics research cluster (HIRC). She was involved in professional teacher development for high school teachers in DKI Jakarta. She is experienced in designing and facilitating teacher professional development training, developing syllabus, designing tasks, developing differentiated instructions, especially in Cambridge IGCSE and A level Biology subject.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Part IV

Effective Teaching and Its Correlates

Part IV Overview

Part IV comprises six chapters. These chapters present studies focusing on various correlates of effective teaching in various international contexts.

Chapter 24 presents a study linking effective teaching behaviour and teachers' intrinsic orientation for the profession (TIOP) in the Netherlands. The study found that the link between effective teaching behaviour and TIOP is mediated by perceived self-efficacy. Background factors including qualification, age, and gender moderate the link between effective teaching behaviour and TIOP. Chapter 25 reports a study about the benefit of online training using reflective teaching and classroom observation measuring effective teaching for improving preservice teachers' reflection in China. Chapter 26 presents a study linking effective teaching and inspiring teaching to student engagement in Hong Kong. It concludes that effective teaching and inspiring teaching are related to student engagement, but differential links between effective and inspiring teaching and student engagement are visible. The study reveals that the dimensions of effective teaching are related to overall teaching quality.

Chapter 27 describes a study linking student perceptions of teaching behaviour and components to learning and motivation in the Norwegian context. The study includes teaching components such as perceived relevance of the content taught, the quality of instruction, the teachers' interest and enthusiasm, and the link between perceived instructional quality and perceived fulfilment of psychological needs. The chapter concludes that students reported lack of intrinsic motivation and experienced low levels of content relevance, and discusses conditions worth investigating when aiming to foster pupils' deep learning and motivation. Chapter 28 presents a study from Singapore examining the influence of teacher beliefs about teaching and learning on students' learning. Specifically, the chapter focuses on understanding how teachers' beliefs affect classroom decisions determining students' learning space and processes in the context of school reform implementation. The chapter provides scenarios illustrating how contextual forces such as curricular content,

national assessments, and achievement-based placement approaches influence teachers' beliefs and practices.

Chapter 29 illustrates how teaching behaviour may differ depending on the perspectives used. The study reported in this chapter compares teachers' perceived and observed effective teaching behaviour in relation to the career phase in the UK. The study found that perceived effective teaching behaviour remains relatively stable throughout teachers' careers; however, their observed effectiveness changes considerably. An increase in teaching effectiveness was observed during middle-phases of teachers' careers, followed by a decrease during the later career phases.

Taken together, these chapters provide insights into effective teaching and its variant or corresponding concepts, in relation to various correlates from a wide range of educational systems. The part highlights the importance of taking into account correlates and contextual forces in studying and improving teaching.

Chapter 24

Teachers' Intrinsic Orientation, Self-Efficacy, Background Characteristics, and Effective Teaching: A Multilevel Moderated Mediation Modeling



Xiangyuan Feng , Michelle Helms-Lorenz , and Ridwan Maulana 

Abstract Teachers' intrinsic orientation for the profession (TIOP) refers to a compound trait derived from the meaningfulness and positive affect teachers attribute to the profession. It can be validly measured by three conceptually correlated yet empirically separable factors of autonomous motivation, enthusiasm for teaching, and enthusiasm for the subject. Grounded in the previous findings of non-significant direct relationships between TIOP and effective teaching, the present study further tested the hypothesized indirect relationships between the two constructs. To better understand the underlying relational mechanisms, the potential mediating role of self-efficacy and the moderating effects of both teacher- and school-level background factors were addressed in single- and multi-level models. A total of 239 beginning teachers from 32 Dutch secondary schools responded to the questionnaires at the beginning of the first career year. Actual teaching behaviour was observed by means of classroom observations. The results of lower-level mediation analysis confirm the mediating effect of self-efficacy on the relationship between TIOP and activating teaching behaviour at career entry. The results of single- and cross-level moderated mediation analysis show that self-efficacy significantly mediates the links between TIOP and three specific teaching behaviour domains: providing safe and stimulating learning climate, classroom management, and clarity of instruction. These effects were respectively moderated by teachers' qualification, age, and gender. The present study makes a unique contribution to understanding the importance of TIOP for beginning teachers' well-being and effective teaching, providing insights for both teacher educators and mentors.

Keywords Teacher intrinsic orientation · Self-efficacy · Background variables · Effective teaching

X. Feng (✉) · M. Helms-Lorenz · R. Maulana
Department of Teacher Education, University of Groningen, Groningen, The Netherlands
e-mail: xiangyuan.feng@rug.nl

© The Author(s) 2023
R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_24

1 Introduction

Teachers' psychological characteristics have long been considered to influence teaching effectiveness (e.g., Barr, 1952). A growing body of literature has highlighted the predictive value of teachers' motivational-affective factors for their teaching quality. Past studies have shown that teachers exhibit more adaptive and operative behaviours at work if they possess higher levels of intrinsically-oriented motivation (e.g., Hein et al., 2012; Hong et al., 2009; Malmberg, 2008; Pelletier et al., 2002; Roth et al., 2007) and positive affect (e.g., Kunter et al., 2008; Moè et al., 2010; Retelsdorf et al., 2010). Based on these findings, Kunter and Holzberger (2014) proposed the compound trait of teachers' intrinsic orientations (TIOs) and extended plausible processes through which TIOs may affect teaching effectiveness. In addition to the direct links, TIOs are claimed to indirectly affect occupational performance via increased classroom effort (de Jesus & Lens, 2005; Feldon, 2007), long-term persistence in professional development (Watt & Richardson, 2008; Lohman, 2006), and well-being (Klusmann et al., 2008). Moreover, Kunter (2013) postulates that these motivational and affective factors may also interact with individual characteristics and situational contexts to determine the types and quality of teaching behaviours.

However, compared to the quantity of empirical studies on the respective role teacher motivation, emotion, or well-being plays in effective teaching, links between the compound construct of TIOs and teaching behaviour are underexplored. To date, only one study was found that explores the influence of teachers' intrinsic orientation for the profession (TIOP), as a compound teacher trait that reflects the general meaningfulness and buoyancy teachers experience from teaching activities and subject matters they teach, on specific and general teaching behaviours (Feng et al., 2021). The results suggested no direct effects, which warrants the necessity for further testing the potential indirect relationships. With this end in view, the present study makes an initial attempt to explore the mediating role of self-efficacy (i.e., teachers' beliefs in their ability to work effectively), one element of teacher well-being (van Horn et al., 2004), in TIOP-teaching behaviour links, by taking into account the specificity of contexts and the hierarchical structure of data. The present study aims to enrich the knowledge base of teacher motivation and teaching effectiveness in two ways. Firstly, the exploration of the indirect TIOP-teaching behaviour links brings new insights into the plausible complex mechanisms underlying the transformation of internal psychological traits into actual teaching behaviour. Secondly, the involvement of multilevel boundary conditions addresses the contextual specificity of TIOP-teaching behaviour link, with regard not only to the relationship strength but also to its direction. Specifically, examining the effects at both lower and higher levels simultaneously may prevent an overestimation of the main effect of teacher-level variables that is typical in hierarchical data.

2 Literature Review

2.1 *Teacher Motivation and Effective Teaching Behaviour*

It has long been acknowledged in educational research that teacher motivation plays a key role in nurturing teaching effectiveness (de Jesus & Lens, 2005; Miller et al., 2008). Studies employing self-determination theory (SDT; Deci & Ryan, 1985, 2000) have established strong associations between teaching practice, student learning, and teacher's autonomous motivation (i.e., deep-rooted or fully internalized endorsement of task value, for example, teachers' believe that teaching is meaningful for self's gratification and students' growth) (for a review, see Slemp et al., 2020). Activated by a full sense of meaningfulness for self and others (Deci et al., 2017; Ryan & Deci, 2017), autonomous motivation is assumed to be associated with higher levels of functional behaviors (Ryan & Deci, 2000). Specifically, Pelletier et al. (2002) identified a positive relationship between Canadian teachers' autonomous motivation and self-reported provision of autonomy support for students. Built upon this finding, Taylor and Ntoumanis (2007) and Taylor et al. (2008) found multiple benefits of autonomous motivation on the use of three motivational strategies (i.e., autonomy support, structure, and involvement) reported by physical education (PE) teachers in the U.K. Similarly, Roth et al. (2007) concluded from their investigation in Israeli elementary schools that teachers' reported autonomous motivation positively predicted student-perceived autonomy-supportive activities, which in turn yielded increased student motivation for learning. Consistent findings were also documented in research across a range of contexts such as Hong Kong secondary schools (Lam et al., 2009), Spanish EFL classrooms (Bernaus et al., 2009), Indonesian junior high schools (Abbas, 2013), and Flemish PE teachers across educational levels (Van den Berghe et al., 2014).

In addition to the consequence of motivational strategies, Hein et al. (2012) also concluded in a cross-national study including Estonia, Hungary, Latvia, Lithuania, and Spain that intrinsically motivated teachers exhibited more student-centered and productive styles of teaching. In the Indonesian secondary school context, teachers' autonomous motivation was positively related to classroom management skills and clarity of instruction (Irnidayanti et al., 2020). In sum, the cumulative evidence reveals a clear relevance of teacher-perceived autonomous motivation with certain aspects of effective teaching. It can be concluded that, in general, teachers who perceive their work as intrinsically worthwhile and meaningful are likely to exhibit higher levels of effective teaching behaviours.

2.2 *Teacher Enthusiasm and Effective Teaching Behaviour*

The topic of teacher enthusiasm in general has captured the interest of educational practitioners and researchers in the past decades for multiple reasons (Keller et al., 2016). Initially characterized in teaching effectiveness research as an indicator of

effective teachers, teaching strategies, and course quality (e.g., Gentry et al., 2011; Moulding, 2010; Walberg & Paik, 2000), teacher enthusiasm manifests itself in a set of outward teacher behaviours perceivable to the observers and students in large scale evaluations. Under a process-product paradigm of this research strand, teacher enthusiasm is characterized by energetic and humorous teaching, sustained student interest (post-hoc analysis without a proactive underlying theory of enthusiasm; e.g., Marsh, 1982, 1994; Marsh & Ware, 1982), student-teacher rapport, and safe and stimulating teaching (Jackson et al., 1999).

Later, Kunter et al. (2008) reconceptualized teacher enthusiasm by shifting the focus of interest from visible “enthusiastic expressiveness” to the relatively hidden “enthusiastic experience” of teachers. Deviating from the cumulative studies on displayed teacher enthusiasm, they proposed the concept of *experienced enthusiasm* and referred to it as “the degree of enjoyment, excitement, and pleasure that teachers typically experience in their professional activities” (Kunter et al., 2008, p. 470). In doing so, these scholars theoretically differentiated the affective and behavioral approaches of teacher enthusiasm and suggested the former as the antecedent to prompt the latter (Frenzel et al., 2009; Kunter et al., 2008, 2011). Furthermore, they recognized two conceptually different, yet correlated sub-dimensions of experienced teacher enthusiasm, one for the subject being taught (i.e., enthusiasm for the subject) and the other for the teaching activity itself (i.e., enthusiasm for teaching) (Kunter et al., 2008, 2011).

The reconceptualization of teacher enthusiasm as an affective trait is also mirrored by the instrument to measure it. Kunter et al. (2008, 2011) put aside the high/low-inference instruments for student perceptions (Frenzel et al., 2009; Patrick et al., 2000; Wheelless et al., 2011) or observer ratings (e.g., Brigham et al., 1992; Natof & Romanczyk, 2009) frequently used in the teaching effectiveness research. Instead, they developed and refined self-reports measures to assess teachers’ experienced enthusiasm in forms of their general impression and evaluation for the enjoyment and pleasure they experience at work (one teaching-specific subscale and one subject-specific subscale). Self-reported enthusiasm for teaching, but not that for the subject, was found to be associated with secondary school teachers’ higher levels of classroom management skills and cognitively activating and supportive teaching, which subsequently benefited students’ motivation and academic achievement (Kunter, 2013; Kunter et al., 2008). In a nutshell, studies generally suggest that teachers who perceive teaching as intrinsically pleasant are more likely to excel in certain teaching behaviour domains.

2.3 Teachers’ Intrinsic Orientation and Effective Teaching Behaviour

Grounded in SDT and teaching effectiveness perspectives, Kunter and Holzberger (2014) encapsulates the conceptually close, yet separable, intrinsic factors of teachers’ orientations into the compound trait TIOs. They refer to TIOs as “the habitual

inter-individual differences between teachers in the degree to which they experience positive emotions and high meaningfulness in their profession” (Kunter & Holzberger, 2014, p. 86). In the theory-led model constructed by Kunter and Holzberger (2014), TIOs is hypothesized as an essential correlate of teacher well-being (e.g., self-efficacy, job satisfaction, burnout) and professional effort at workplace (e.g., engagement and persistence in professional learning, classroom efforts). These teacher factors in turn benefit instructional quality and subsequent student outcomes. More specifically, it is assumed that the positive influence of TIOs on effective teaching behaviour can be explained by both direct psychology-behavior links and indirect relationships mediated by teachers’ situational classroom effort (de Jesus & Lens, 2005; Feldon, 2007), well-being (Klusmann et al., 2008), and long-term persistence in professional development (Watt & Richardson, 2008; Lohman, 2006) (see Fig. 24.1). Additionally, teachers’ motivational and affective traits are postulated to interact with individual characteristics and situational contexts to determine the types and quality of teaching behaviour (Kunter, 2013). The innovative value of this model lies in its additional explanation for the underlying process where various psychological and behavioral traits of teachers interplay for better functioning across contexts.

IN LIGHT OF KUNTER AND Holzberger (2014)’s theory, TIOs have been further crystallized by being rephrased into teachers’ intrinsic orientation for the profession (TIOP) (Feng et al., 2021). The construct validity of TIOP was empirically tested in terms of its dimensionality via teachers’ self-reported autonomous motivation (i.e., a cognitive-evaluative factor reflecting the meaningfulness teachers ascribe to the profession) and experienced enthusiasm for teaching and for the subject (i.e., affective-evaluative factors to elicit teachers’ positive emotional experience) (Feng et al., 2021) (see Fig. 24.1). The results concluded that TIOP can be constructed as a compound trait of teachers with three subdimensions. However, the empirical testing of TIOP’s predictive validity for the quality of the general as well as specific observed teaching behavior (i.e., providing safe and stimulating learning environment, classroom management, clarity of instruction, intensive and activating

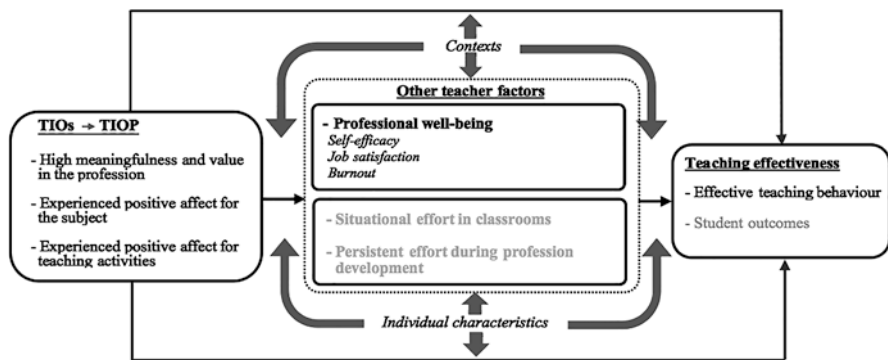


Fig. 24.1 A model of the relationships between TIOs/TIOP and effective teaching adapted from Kunter and Holzberger’s (2014) theory (concepts not included in the present study are blurred)

teaching, differentiated instruction, teaching learning strategies) indicated that there was no significant direct relationship between TIOP and the six teaching behaviour domains (Feng et al., 2021). Consequently, it is hypothesized that the effects of TIOP on displayed teaching behaviors may be indirect and may potentially be influenced by certain teacher characteristics in different boundary conditions.

In other words, instead of functioning as a direct facilitator, TIOP may indirectly profit the quality of displayed teaching behaviour through its positive effects on teachers' psychological well-being and the subsequent intentional efforts they invest in the profession. However, the strength of these direct and indirect effects may vary across teachers with different personal characteristics (e.g., age, gender, academic qualification) and working contexts (e.g., class size, school culture, principal leadership). Unfolding such complex interplay of teacher factors is therefore considered of great value to understand the process of successfully transforming (student) teachers' inner power into the actual profits for themselves (i.e., well-being and professionalization) and the students (i.e., teaching and learning effectiveness). As an initial step of this exploration, the present study examines the role of teacher self-efficacy in mediating the links between TIOP and different domains of teaching behavior, while considering the influence of the relevant background factors at teacher and school levels.

2.4 Mediators and Moderators of the Relationship Between TIOP and Effective Teaching Behaviour

2.4.1 Self-Efficacy as a Mediator

As illustrated in Fig. 24.1, TIOP can function as either a direct or an indirect resource for instructional quality through teachers' increased well-being. While a rigorous analysis of all possible mediators is beyond the scope of the present study, the focus is on the mediating role of self-efficacy as a representative factor of teachers' occupational well-being (Van Horn et al., 2004). Since the concept of TIOP is relatively novel and empirical research on TIOP is scarce, existing literature on the sequential connections among TIOP-related concepts such as autonomous motivation, enthusiasm, self-efficacy, and teaching behaviors are elaborated for reference.

A large body of empirical literature has documented the benefits of autonomous motivation for teachers' psychological well-being and functioning in diverse contexts (e.g., Fernet et al., 2017; Ryan & Deci, 2000, 2017). More specifically, autonomously motivated teachers are more likely to experience higher levels of self-efficacy (Gagné et al., 2015), sense of accomplishment (Ryan & Deci, 2000), job satisfaction (Collie et al., 2016), and overall satisfaction of life (Pauli et al., 2018). In addition to intrinsically-orientated motivation, experienced enthusiasm also bears a close link to enhanced well-being of teachers (Keller et al., 2016). Enthusiastic teachers were found to be more self-efficacious (Kunter et al., 2011), and satisfied with their work and life (Kunter, 2013; Kunter et al., 2008, 2011). In sum,

autonomous motivation and experienced enthusiasm of teachers seems to go hand in hand with self-efficacy and other factors of well-being.

Self-efficacy as a primary indicator of teachers' well-being has been both theoretically and empirically supported to predict teachers' beliefs about instructional behaviors (Ross, 1992; Skaalvik & Skaalvik, 2007; Tschannen-Moran & Hoy, 2001). Teachers with lower levels of self-efficacy are more likely to experience setbacks in teaching (Betoret, 2006). A meta-analysis of 43 self-efficacy studies done by Klassen and Tze (2014) reveals a significant medium effect size ($r = .28$) of self-efficacy on evaluated teaching performance (via principal, supervisor, student ratings), which is consistent with the prior self-efficacy studies outside the education discipline (e.g., $r = .38$; Stajkovic & Luthans, 1998). This conclusion was further clarified in another review study (Zee & Koomen, 2016) which identified the consequence of in-service teacher self-efficacy on teaching behaviors such as process-oriented instruction and differentiation, activating teaching strategies, inclusive practices and referral decisions, classroom management skills (both instructional and behavioral), classroom goal structures, and emotional support. In sum, more efficacious teachers are likely to exhibit a learner-centered constructivist style of teaching (Temiz & Topcu, 2013). However, the role of self-efficacy as a mediator of teacher motivation and teaching behaviour is unclear.

2.4.2 Teacher Characteristics and Contexts as Moderators

Moderators are considered very informative in social science research since they underline the boundary conditions of a theory's generalizability (Whetten, 1989). Informed by the empirical evidence on how certain contextual and personal factors influence teaching behaviors, the present study aims to test the contextualized relationship between self-efficacy and teaching behaviour. Considering that teachers are naturally embedded in hierarchical school structures, the contextual factors that may impact their professional practices should be considered in a multi-level design (e.g., school, classroom, teacher). The regional or school level factors such as the dynamics and size of student population, the student-teacher (employment size) ratio, financial distribution for school management and teacher professionalism may influence the attraction, retention, and growth of high-quality beginning teachers (for a review, see van der Pers & Helms-Lorenz, 2019). Specifically in the context of Dutch secondary schools, 11% to 22% of the variance in beginning teachers' observed teaching behaviour was attributed to school-level characteristics (van der Pers & Helms-Lorenz, 2019). Among them, effects of urbanization degree and student population decline were found on stimulating teaching, classroom management, and adaptive instruction, respectively. Furthermore, many schools provide novices and veterans with different degrees of learning opportunities and infrastructures. For instance, professional development schools (PDSs) in the Netherlands collaborate with education institutes to support teachers by means of sustainable and collaborative activities, which in turn fosters beginning teachers' general teaching behaviour during their first career year (Helms-Lorenz et al., 2018).

Apart from the contextual factors at higher levels, personal characteristics at the teacher level such as *gender* (e.g., Opdenakker et al., 2012; Opdenakker & Van Damme, 2007; Van Petegem et al., 2007), *age* and *teaching experience* (e.g., Kini & Podolsky, 2016; Ladd & Sorensen, 2015; Maulana et al., 2015), *educational background* and *certification* (see Tatto et al., 2012; van der Pers & Helms-Lorenz, 2019) are, in varying degrees, related to teachers' instructional quality. Amongst these factors, cumulative training and practical experience predominantly avail teachers improved instructional skills (e.g., van der Pers & Helms-Lorenz, 2019), and male teachers are found to exhibit better instructional (Maulana et al., 2015) and relational skills (e.g., classroom management, student interaction, cooperativeness) (e.g., Opdenakker et al., 2012). Furthermore, since the process of teaching and learning is inherently interactive and reciprocal, *student factors* (at class, school, regional levels) have been revealed to affect teachers' professional well-being and teaching effectiveness (Kunter & Holzberger, 2014). For example, schools with a predominant proportion of low socioeconomic-status (SES) students were found to hinder beginning teachers' workplace learning (Ronfeldt, 2012) and inhibit peer/colleague cooperation (Opdenakker & Van Damme, 2007). Comparatively, smaller classes may engender more individualized teaching and teacher-student interaction, after controlling for prior pupil attainment, gender, and special education needs (Blatchford et al., 2011). Nevertheless, the moderating effect of certain personal and contextual background factors on the link between teacher motivation and teaching behaviour requires further investigation.

3 The Current Study

Whereas novices in most occupations generally begin with minor duties and progressively receive more challenging assignments along their trajectory of professionalization, beginning teachers tend to receive full pedagogical and organization responsibilities immediately after career entry (Tynjälä & Heikkinen, 2011). Increasingly strained by instructional challenges (e.g., heavy workload, students' low engagement and misbehavior, differentiated teaching) and a discrepancy between professional efficacy and preparedness, beginning teachers experience prevalent praxis shock (Ashby et al., 2008; Hoy & Spero, 2005). This problem seems to subsequently jeopardize professional well-being and motivation, leading to rising teacher attrition and shortages in the longer term (e.g., Helms-Lorenz et al., 2016). In view of such concerns, the present study assigned research priority to the assessment of beginning teachers' TIOP and delved into the relationships between teachers' self-perception (i.e., TIOP and self-efficacy) and preparedness (i.e., general and specific teaching behaviour) at career entry.

Since the strengths of these relationships might vary across contexts (Blömeke et al., 2016), no prior assumption was made regarding the moderating effects of one

particular background variable on the link between TIOP and teaching behaviour. Instead, a general hypothesis was developed only on the existence of personal or contextual factors as moderators in the efficacy-teaching behavior link. By employing an exploratory approach, the influence of TIOP on the specific and general teaching behaviour via self-efficacy were scrutinized for its context-(in)dependency. To achieve this purpose, the following research questions were to be answered:

- 1 Does teachers' self-efficacy mediate the relationships between TIOP and the specific and general observed teaching behaviour?
- 2 Do teachers' demographics (i.e., age, gender, education degree, qualification types) and contextual characteristics at teacher-level (i.e., class size, students' gender, age, prior academic scores) moderate the mediating effect of self-efficacy in the relationship between TIOP and teaching behaviour?
- 3 Do school characteristics (i.e., school size, school type, student teacher ratio, employment size, gender and age distribution of teacher population; student SES) moderate the mediating effect of self-efficacy in the relationship between TIOP and teaching behavior?

4 Methods

4.1 *Participants and Procedure*

The present study was a part of a 3-year research project on the teacher induction program implemented in the northern Netherlands (in Dutch: *Inductie in het Noorden (INO)*), which was subsidized by the Dutch government. After the research objectives and protocols were developed, 239 beginning teachers ($N_{\text{female}} = 144$, $M_{\text{age}} = 28.74$), ranging from 21 to 61 years of age and of all subject matters, voluntarily participated in the project at career entry. They were unevenly distributed among 32 Dutch secondary schools ($N_{\text{teachers per school}} = 1-21$). Specifically, three cohorts of teachers were included. Cohort 1 ($N = 73$) were surveyed with the questionnaires of TIOP and self-efficacy between November and December in 2014, cohort 2 ($N = 78$) between October and November in 2015, and cohort 3 ($N = 88$) between October and November in 2016. In addition to self-reports, beginning teachers were observed by well-trained observers and rated on the quality of the six domains of teaching behavior displayed in the classroom. The Dutch version of these instruments was employed in this study after translation and back translation procedure was conducted (Hambleton, 1994). School contextual factors and personal characteristics were collected from secondary sources or public databases. In order to increase response rates, teachers who participated throughout the INO project were provided with a €30 gift voucher and annual feedback.

4.2 Measures

TIOP. Dutch beginning teachers' TIOP was measured using a validated TIOP scale, which consists of the sub-dimensions of experienced enthusiasm for teaching (4 items), experienced enthusiasm for subject (4 items), and autonomous motivation (3 items) (Feng et al., 2021) (see Appendix Table 24.A1). Teachers' responses were scored using four-point Likert scales ranging from 1 (completely disagree) to 4 (completely/strongly agree). Considering the multidimensional second-order structure of TIOP, omega (0.91, 0.92) and omega hierarchical (0.79, 0.78), instead of alpha, were selected as the reliability coefficients. The estimates of omega (hierarchical) indicated that the total score of the compound TIOP scale primarily reflects the characteristics of the general factor TIOP while also leaving space to capture the specificity of sub-factors in the lower order constructs. However, the low internal consistency of the autonomous motivation subscale (alpha = .436) is most probably due to the limited number and heterogeneity of items (see Appendix Table 24.A1). This finding suggests that this subscale be better used as part of the TIOP measure rather than an independent scale.

Self-efficacy. We used the Teachers' Sense of Efficacy Scales (TSES; Tschannen-Moran & Hoy, 2001) to measure teachers' perceived self-efficacy (see Appendix Table 24.A1). Consisting of 24 items, the scale covers three domains of teacher efficacy: efficacy for instruction (8 items), efficacy for classroom management (8 items) and efficacy for student engagement (8 items) (see Appendix Table 24.A1). Teachers responded on a five-point Likert scale, ranging from 1 (nothing) to 5 (a great deal). Acceptable to high reliability coefficients of alpha (0.62–0.94) of both the general and sub-scales were reported across contexts and over time (Duffin et al., 2012; Feng et al., 2021; Helms-Lorenz et al., 2018; Tschannen-Moran & Hoy, 2001). In the present study, TSES was employed to measure beginning teachers' general teaching self-efficacy. In addition, raw scores rated on the 5-point scale were converted to 4-point scale, using the linear transformation equation: $(\text{Max}_{\text{new}} - \text{Min}_{\text{new}}) \times (X - \text{Min}_{\text{old}}) / (\text{Max}_{\text{old}} - \text{Min}_{\text{old}}) + \text{Min}_{\text{new}}$.

Observed teaching behaviors. Six domains of observable teaching behavior (i.e., providing safe and stimulating learning environment, classroom management, clarity of instruction, intensive and activating teaching, differentiated instruction, teaching learning strategies) were assessed by well-trained observers using the validated Dutch version of International Comparative Analysis of Learning and Teaching (ICALT) instrument (Maulana et al., 2017; Van de Grift et al., 2014). The instrument consists of 120 low-inferential items specifying observable teaching behaviours, which are categorized into 32 high-inferential items as indicators of the aforementioned six behavioral domains. Each indicator was rated on a four-response category (1 = "mostly weak, 4 = "mostly strong"). These generic behavioral domains have been identified as essential for supporting and maximizing students' learning, thus reliably manifesting the effectiveness of teaching in classrooms. The validity and reliability of the measure have been proven good across various national contexts (alpha from 0.74 to 0.92) (Maulana et al., 2017, 2020).

Background variables. The multilevel background factors included in this study are teachers' demographics (i.e., age, gender, education degree, and qualification types) and contextual characteristics at teacher-level (i.e., class size, students' gender, age, and prior academic scores), and school-level (i.e., school size, school type, student teacher ratio, employment size, gender and age distribution of teacher population; student SES). Among them, teacher and class characteristics were recorded together with the questionnaires on teaching behaviour or the supervision monitor. Professional development school status (VORaad), school sizes (DUO, 2015, 2016; VOION, Arbeidsmarkt en Opleidingsfonds Voortgezet Onderwijs, 2016), and SES of neighbourhoods (Sociaal en Cultureel Planbureau [SCP], 2014) are all secondary data from mentioned sources. These background factors were included in the models as moderators of the relationships between self-efficacy and teaching behaviour.

4.3 Data Analysis

4.3.1 Preliminary Analysis

The proportion, patterns, and mechanisms of data missingness were scrutinized for the sake of unbiased estimates of parameters, statistical power, and generalizability of findings (Dong & Peng, 2013). Initial analysis results indicate a missing rate of 0% to 16.3% on key variables (i.e., TIOP-related factors, self-efficacy, observed teaching behaviors) (see Table 24.1). Although about 15%–20% data missingness is common in educational and psychological studies (Enders, 2003), missingness above 10% is considered consequential to statistical inferences (Bennett, 2001). Therefore, all key variables were further assessed in terms of the mechanisms of missingness using Little's Test of Missing Completely at Random (MCAR) (Little, 1988).

Table 24.1 Proportion of missingness in all variables

<i>Focal factors</i>	Valid N	Missing
TIOP-related factors	239	0
Self-efficacy	239	0
Teaching behaviors	200	39 (16.3%)
<i>Background factors</i>		
Age	238	1 (0.4%)
Gender	239	0 (0.4%)
Qualification type	237	2 (0.8%)
Degree type	236	3 (1.3%)
Class size	193	46 (19.2%)
Student mean age	193	46 (19.2%)
Student gender distribution	129	110 (46.0%)
Student prior score	185	54 (22.6%)
School contexts	168–237	2–71 (0.4–29.7%)

The construct validity of focal latent variables (i.e., second-order TIOP, second-order self-efficacy, correlated teaching behavior domains) was subject to confirmatory factor analyses (CFAs) using Mplus 8.3, on condition that the plausibility of MCAR or MAR was justified in the evaluation of cross-sectional missingness. Factors scores were thereby calculated and used for the following structural equation modeling (SEM). Furthermore, by modeling TIOP and self-efficacy in the same model (with their correlation set free), the average variance extracted (AVE) and composite reliability (CR) were estimated so as to examine the discriminant and convergent validity of the individual-level self-report data (see Appendix Table 24.A2).

4.3.2 Single and Multilevel Mediation Analysis

To test the mediating effects of self-efficacy, simple mediation models were first constructed, where the quality of general and specific teaching behavior was regressed on TIOP via self-efficacy. Goodness-of-fit indices were estimated. Preacher et al.'s (2010) Monte Carlo bootstrap method was applied to generate 95% confidence intervals (IC) that assists in making conclusions on the significance of the indirect effects. Then, on condition that the rationality of performing multilevel mediation analysis was justified through the intra-class correlations (ICC) of teaching behavior domains (ICCs = [0.100, 0.178]), lower level mediation models were constructed (see Fig. 24.2). In these random effect models, all causal paths were

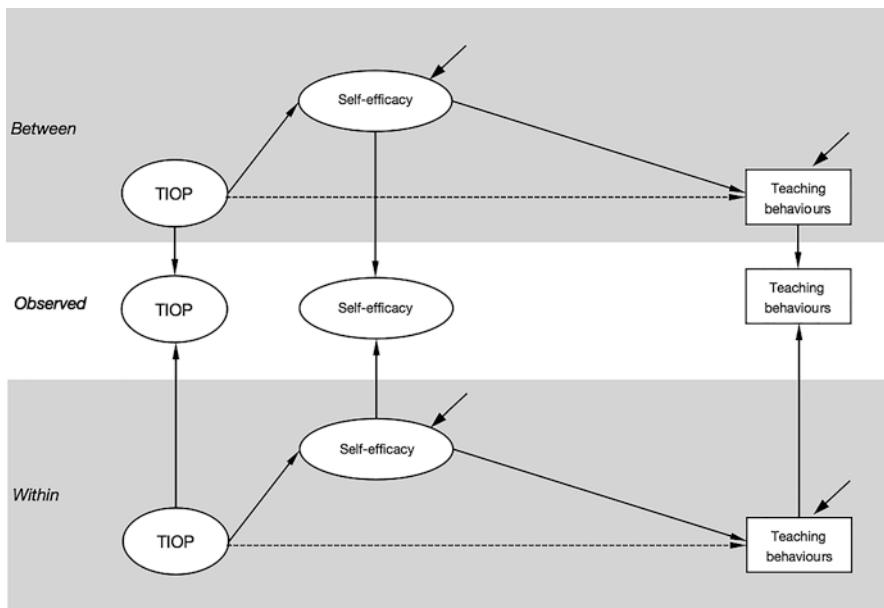


Fig. 24.2 Lower level mediation model between TIOP, self-efficacy, and teaching behavior

allowed to vary between school units. We compared their related fit indices of Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and sample size adjusted BIC (ssaBIC) and then estimated the multilevel mediating effects (see Preacher et al., 2011; teacher-level mediation = $a_{L1} \times b_{L1}$ + L2 covariance of a_{L1} and b_{L1} ; school-level mediation = $(a_{L1} + a_{L2}) \times (b_{L1} + b_{L2})$). Markov chain Monte Carlo (MCMC) estimation method was applied to assess the significance of school-level mediation at 95% IC. However, given that there are 7 clusters with only one member, these clusters contribute to the estimation of school-level parameters rather than individual-level ones, resulting in less individual-level power. Consequently, path estimates and confidence intervals calculated by MCMC were only reported for school-level mediation effects.

4.3.3 Single and Multilevel Moderated Mediation Analysis

After the testing of simple and lower level mediation models, background factors from two levels (i.e., teacher, school) were added to the model (see Fig. 24.3). It is presumed that school contextual characteristics are identical and thus function in a uniform manner towards individual teachers in the same schools. Therefore, a set of simple and cross-level models (i.e., teacher-school levels) were formulated, in which independent (i.e., TIOP) and dependent variables (i.e., specific and general teaching behaviors), mediator (i.e., self-efficacy), and teacher characteristics are level 1 (L1) variables, whereas school contexts are level 2 (L2) variables (see Fig. 24.3). Due to the limited sample size, the moderating effect of each factor was explored successively. The software Mplus 8.3 was used since it allows the examination of mediation and moderation in one single model and enables correct estimation of parameters and errors. In these models, the effects of L2 moderators were specified as random.

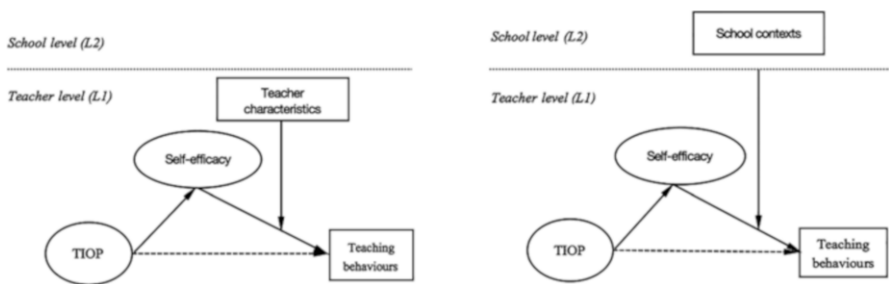


Fig. 24.3 Successive mediation models moderated by background or contextual factors at different levels

5 Results

5.1 Preliminary Results

Preliminary analysis was conducted to examine the possible consequences of missing values in data and to test the measurement validity of established instruments in the target context. The results of Little's tests ($X^2 = 28.271$, $df = 35$, $p = .783$) suggest that missing values on key variables (i.e., TIOP-related factors, observed teaching behaviors) were randomly distributed and did not depend on any other measured or non-measured variable (Graham, 2009). Consequently, cases can be dropped listwise or pairwise during factor analysis and SEM, and implementation of the maximum likelihood approach for handling missingness is also supported. The descriptive statistics of the raw scores of self-reports, observation, and teacher characteristics, along with their bivariate correlations are shown in Table 24.2, with the scale scores of self-efficacy converted to 4-point scaling. The reliability coefficients of alpha for each sub-scale were also estimated.

Based on the above findings, CFAs of teacher-level observation and self-reports were legitimate, which yielded good model fits: (1) $X^2_{\text{TIOP}} (41, N = 239) = 83.841$, $CFI_{\text{TIOP}} = 0.986$, $TLI_{\text{TIOP}} = 0.981$, $RMSEA_{\text{TIOP}} = 0.066$, and $SRMR_{\text{TIOP}} = 0.055$, $\lambda_S = [0.446, 0.949]$; (2) $X^2_{\text{SE}} (0, N = 239) = 0.000$, $CFI_{\text{SE}} = 1.000$, $TLI_{\text{SE}} = 1.000$, $RMSEA_{\text{SE}} = 0.000$, and $SRMR_{\text{SE}} = 0.000$, $\lambda_S = [0.620, 0.829]$; (3) $X^2_{\text{TB}} (0, N = 200) = 0.000$, $CFI_{\text{TB}} = 1.000$, $TLI_{\text{TB}} = 1.000$, $RMSEA_{\text{TB}} = 0.000$, and $SRMR_{\text{TB}} = 0.000$, $r = [0.113, 0.706]$. In general, all item loadings and factor correlations are significant and range from moderate to high, except the link between stimulating teaching and teaching learning strategies ($r = 0.113$, $p = 0.086$). The calculation of factor scores instead of means was warranted due to the heterogeneous loadings among three sub-domains of TIOP ($\lambda_{\text{ET}} = 0.949$, $\lambda_{\text{ES}} = 0.812$, $\lambda_{\text{AM}} = 0.446$, $ps < .001$) and self-efficacy ($\lambda_{\text{SE1}} = 0.829$, $\lambda_{\text{SE2}} = 0.620$, $\lambda_{\text{SE3}} = 0.753$, $ps < .001$).

To examine the discriminant and convergent validity of the teacher-level self-report data, TIOP and self-efficacy were estimated in a single model (see Appendix Fig. 24.A1). Goodness-of-fit indices indicated good fit, $X^2 (73, N = 239) = 124.250$, $CFI = 0.983$, $TLI = 0.979$, $RMSEA = 0.054$, and $SRMR = 0.056$. Based on the reported standardized factor loadings and residual variances, AVEs and CRs were calculated, showing acceptable to satisfactory results ($AVE_{\text{TIOP}} = 0.59$; $CR_{\text{TIOP}} = 0.80$; $AVE_{\text{SE}} = 0.54$; $CR_{\text{SE}} = 0.78$). Since the AVE values of the higher-order TIOP and multidimensional self-efficacy are above 0.5 and those of CR above 0.7, convergent validity was supported (Fornell & Larcker, 1981). Besides, given that the amount of the variance captured by TIOP or self-efficacy ($\sqrt{AVE} = 0.74\text{--}0.77$) were greater than their correlation ($r = 0.613$), discriminant validity was supported (Fornell & Larcker, 1981). In general, the CFA results prove that the established instruments applied in this study are valid measures of beginning teachers' TIOP, self-efficacy, and teaching behavior, respectively, in the Dutch context.

Table 24.2 Descriptive statistics, bivariate correlations, and reliability coefficients of alpha

	Mean (SD)	Skewness	Kurtosis	Correlation								alpha								
				F1	F2	F3	F4	F5 ₁	F5 ₂	F5 ₃	F5 ₄		F5 ₅	F5 ₆	F6	F7	F8			
<i>Focal factors</i>																				
F1: Enthusiasm for teaching	3.53 (0.40)	-0.25	-1.29																	.761
F2: Enthusiasm for subject	3.30 (0.49)	-0.09	-0.59	.586*																.843
F3: Autonomous motivation	2.86 (0.51)	-0.32	0.44	.179**	.192**															.436
F4: Self-efficacy	2.73 (0.29)	0.40	1.15	.421**	.380**	.189**														.898
F5 ₁ : Stimulating teaching	3.38 (0.52)	-1.01	2.17	.111	.000	.033	.060													.809
F5 ₂ : Classroom management	3.16 (0.58)	-0.51	0.08	.136	.078	-0.066	.073	.549**												
F5 ₃ : Instruction clarity	3.00 (0.55)	-0.18	-0.60	.073	.038	-0.053	.068	.597**	.642**											
F5 ₄ : Activating teaching	2.49 (0.59)	-0.15	-0.24	.096	.069	.014	.120	.469**	.478**	.706**										
F5 ₅ : Differentiated teaching	1.81 (0.68)	0.68	-0.11	.063	.089	.006	.107	.162*	.318**	.276**	.508**									
F5 ₆ : Teaching learning strat	1.81 (0.69)	0.64	-0.23	-0.016	.008	.004	.083	.113	.202**	.294**	.585**	.494**								

(continued)

Table 24.2 (continued)

Background factors	Mean (SD)	Skewness	Kurtosis	Correlation										alpha		
F6: Age	28.74 (7.61)	2.06	3.76	.019	.123	-.126	.053	-.258**	-.182*	-.162*	-.142*	-.076	-.065			
F7: Gender (male = 1; female = 2)	1.60 (0.49)	-0.42	-1.84	-.031	-.075	.086	-.142*	.000	.054	.001	.036	.125	.065	-.142*		
F8: Qualification type	1.90 (0.36)	-1.24	3.52	-.043	-.043	-.180**	-.007	-.035	.035	.029	.054	.067	.074	-.007	.048	
F9: Degree type	2.19 (2.00)	2.39	4.36	.076	.081	.171**	.070	-.029	-.076	-.077	-.113	-.052	-.159*	.077	-.053	-.823**

* $p < .05$, ** $p < .001$

5.2 *Self-Efficacy as the Mediator*

The analysis of within- and cross-cluster mediation examines the multi-level relationship between TIOP and effective teaching that is mediated by self-efficacy (research question 1: single- and lower-level mediation models). Firstly, every simple mediation model showed acceptable model fit (CFI > .977; TLI > .942; RMSEA < .080; SRMR < .038) (see Appendix Table 24.A2). However, no significant mediating effect of self-efficacy was found on the relationship between TIOP and teaching behaviour. Secondly, all lower level mediation models except TIOP-activating teaching (TB4) showed non-significant indirect effects (unstandardized $\beta_{\text{mediation}_{\text{TB4}}} = -2.300, p = .065; IC_{\text{MCMC}} = [-5.19, -0.23]$) (see Appendix Table 24.A2). In this model, TIOP significantly predicted self-efficacy (unstandardized $\beta = -1.447, p = .023, IC = [-2.696, -0.199]$), which, in turn, predicted TB4 (unstandardized $\beta = 1.589, p < .001, IC = [1.129, 2.049]$). After controlling for the mediator, TIOP was regressed on TB4 with unstandardized $\beta = 16.745 (p < .001, IC = [16.157, 17.333])$. Combining the direct and indirect effects results in a positive and significant total effect (unstandardized $\beta_{\text{total}} = 14.446, p < .001$).

Compared to the non-significant positive mediation (unstandardized $\beta_{\text{mediation}_{\text{TB4}}} = 0.204, p > .05, IC = [-0.112, 0.529]$) in the corresponding single-level model, self-efficacy's mediating effect was negative and significant in the lower level model. This is caused by the stronger between-school links of TIOP-efficacy (unstandardized $\beta_{\text{TIOP-SE}} = -2.315, p < .001; IC = [-3.443, -1.187]$) and of efficacy-TB4 (unstandardized $\beta_{\text{SE-TB4}} = 1.286, p < .001; IC = [1.101, 1.471]$), as illustrated in Fig. 24.4. In the same vein, the direct effect of TIOP on TB4 turned significant in the lower level model due to the stronger between-school effect (unstandardized $\beta_{\text{TIOP-TB4}} = 17.025, p < .001; IC = [16.882, 17.167]$). In general, self-efficacy seemed to partially suppress the effect of TIOP on the quality of intensive and activating teaching at the outset of teaching career. However, such effect is mainly caused by between-school differences, leaving the teacher-level direct and indirect links not statistically significant.

5.3 *Background Variables as Moderators*

The analysis then moved to the estimation of moderated mediation. Whether the mediation effects of self-efficacy were strengthened or weakened by personal (research question 2: single-level models) and school characteristics (research question 3: cross-level models) was examined. In total, four single-level models but no cross-level models were found with significant moderated mediation (see Table 24.3). All models showed similar related fit indices when compared to simple mediation models ($\Delta_{\text{AICs}} = [-14.555, -1.006], \Delta_{\text{BICs}} = [-7.736, 5.946], \Delta_{\text{ssa}_{\text{ICs}}} = [-14.074, 0.394]$). As illustrated in Fig. 24.5, teachers' TIOP positively predicted self-efficacy with $\beta = [0.515, 0.523], p < .05$, while self-efficacy in turn (1)

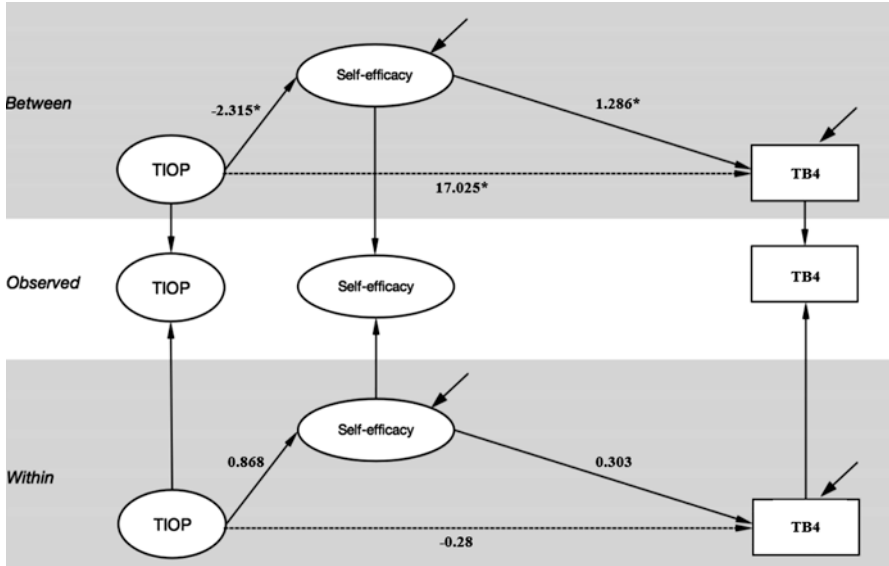


Fig. 24.4 Low-level model with the significant mediating effect of self-efficacy
 * $p < .05$

Table 24.3 Fit indices of simple mediation models

Model	Related fit indices			Mediation		Moderated mediation	
	AIC	BIC	ssaBIC	p	IC_{MCMC}	p	IC_{MCMC}
Model 1_qualification_TB1	996.510	1051.999	1001.285	.012	[-1.203, -0.172]	.003	[0.149, 0.678]
Model 2_qualification_TB2	1042.156	1097.645	1046.931	.036	[-1.630, -0.113]	.034	[0.060, 0.815]
Model 3_age_TB2	1038.156	1093.712	1042.997	.017	[-3.000, -0.365]	.016	[0.013, 0.103]
Model 4_gender_TB3	1034.776	1090.399	1039.683	.043	[0.093, 2.169]	.027	[-1.236, -0.108]

negatively predicted stimulating teaching (TB1) and classroom management (TB2) ($\beta = [-0.843, -0.404], p < .05$), with the latter slopes positively predicted by the interference moderators of qualification or age ($\beta_{interaction} = [0.168, 0.222], p < .05$), or (2) positively predicted clarity of instruction (TB3) ($\beta = [0.604, 0.796], p < .05$), with negative interference moderators of gender or age ($\beta_{interaction} = [-0.194, -0.178], p < .05$).

Specifically, after involving the hypothesized mediators and moderators, the influence of TIOP on TB1 and TB2 was fully suppressed by self-efficacy ($\beta_{mediation_model1} = -0.657; p = .012; IC_{MCMC} = [-1.203, -0.172]; \beta_{mediation_model2} = -0.807; p = .036; IC_{MCMC} = [-1.630, -0.113]; \beta_{mediation_model3} = -1.582; p = .017$;

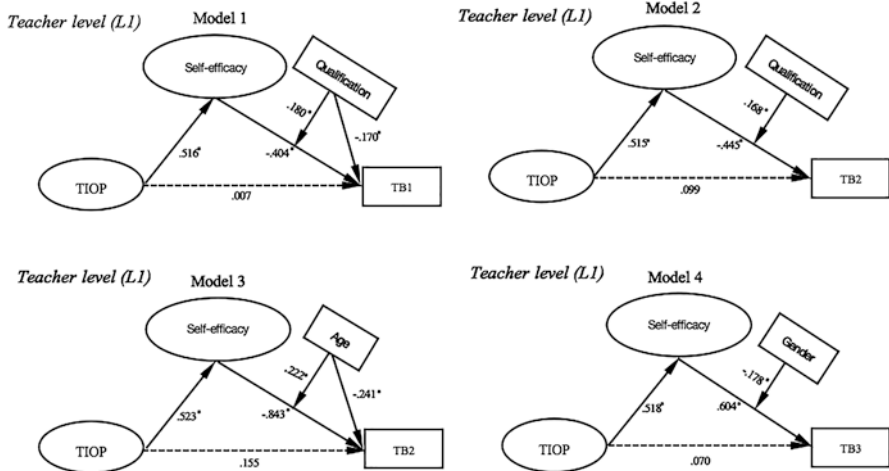


Fig. 24.5 Models with significant effects of moderated mediation * $p < .05$.

$IC_{MCMC} = [-3.000, -0.365]$). The suppression effects on TB1 decrease with qualification ($\beta_{\text{moderatedmediation_model1}} = 0.394; p = .003; IC_{MCMC} = [0.149, 0.678]$). The effects on TB2 also decrease with qualification ($\beta_{\text{moderatedmediation_model2}} = 0.409; p = .034; IC_{MCMC} = [0.060, 0.815]$) and age ($\beta_{\text{moderatedmediation_model3}} = 0.055; p = .016; IC_{MCMC} = [0.013, 0.103]$). Comparatively, self-efficacy was also found to fully mediate the positive effects of TIOP on TB3 ($\beta_{\text{mediation_model4}} = 1.061, p = .043; IC_{MCMC} = [0.093, 2.169]$), and this mediating effect was stronger for males ($\beta_{\text{moderatedmediation_model4}} = -0.639; p = .027; IC_{MCMCs} = [-1.236, -0.108]$). In general, teacher characteristics such as qualification, age, and gender, rather than contextual factors at both teacher and school levels, significantly moderate the indirect links between TIOP and relatively basic and teacher-centered teaching behavior.

6 Discussion and Conclusion

The main purpose of this study was to test the indirect links between TIOP and teaching behaviour built upon the previous work of Kunter and Holzberger (2014). Since the concept of TIOP is relatively novel and relevant empirical research is scarce, the knowledge base of TIOP is still in development. The present study is one of the first to address the theoretical and empirical implication of TIOP, as a compound teacher trait, in teaching effectiveness research.

The first research question was: *Does teachers' self-efficacy mediate the relationships between TIOP and the specific and general observed teaching behaviour?* The findings of simple and lower-level mediation analysis answered this question by providing such evidence that, after considering the naturally nested structure of

teacher workforce, self-efficacy is found to partially suppress the positive relationship between TIOP and activating teaching at the outset of teaching career. This is in line with the findings of Ryan and Deci (2000, 2017), Kunter (2013), and Kunter et al. (2008) about teachers' positive psychological factors (i.e., TIOs, well-being) benefiting effective teaching behaviour, whereas partly inconsistent with Kunter and Holzberger's (2014) hypothesis on self-efficacy as a facilitating mediator. A closer look at the relationships at both levels reveals that self-efficacy does serve as a facilitator at the teacher level, which confirms the empirical findings of Gagné et al. (2015), Kunter et al. (2011), Klassen and Tze (2014), and Zee and Koomen (2016). However, the stronger suppressing effect of self-efficacy found at the school level, caused by the negative TIOP-efficacy link, completely overwhelmed the aforementioned teacher-level effect. Most likely, it is caused by the external school-level factors which have not been internalized by beginning teachers, such as recruitment policies to attract and retain teachers with qualities that are aligned to the school culture.

It seems that the school-teacher mutual selection somehow leads to the gathering of teachers with a discrepancy between TIOP and self-efficacy. One possible explanation of this could be some schools' tendency to attract and recruit enthusiastic teachers who are experiencing praxis shock. Beginning teachers who rate themselves high on TIOP-related scales are more likely to hold higher expectations towards the teaching profession (Ashby et al., 2008) and sometimes more vulnerable to role shock and disillusion. As a consequence, these intrinsically motivated beginning teachers may possess better activating teaching skills to maximize learning outcomes but their actual performance is slightly interfered by the loss of self-confidence in implementing them in classrooms. Comparatively, some other schools may find a majority of their beginning teachers with relatively lower enthusiasm or intrinsic motives yet higher self-efficacy. In their cases, self-efficacy can serve as a buffer to offset the influence of low TIOP on activating teaching skills.

Considering that the strengths of TIOP-efficacy-behavior links might vary across different boundary conditions, the second and third questions were raised: *Do teacher characteristics and school contexts moderate the mediating effect of self-efficacy in the relationship between TIOP and teaching behavior?* Results of single-level moderated mediations answered the second research question, suggesting that personal factors such as qualification, age, and gender significantly moderate certain indirect TIOP-teaching behavior links. However, cross-level model results do not provide any empirical evidence for the moderating effects of school-level characteristics. As a complement of the first conclusion that self-efficacy partially mediates the TIOP-activating teaching link at the school level, moderated mediation results reveal that self-efficacy also fully mediates the relationships between TIOP and three other teaching behaviours (i.e., providing safe and stimulating learning environment, classroom management, clarity of instruction) at the teacher level. Such findings provide further evidence supporting the positive links between TIOs and teacher well-being (e.g., Gagné et al., 2015; Kunter et al., 2011) as well as the

gender effect (e.g., Maulana et al., 2015; Opdenakker et al., 2012) and benefits of teacher experience on effective teaching (e.g., van der Pers & Helms-Lorenz, 2019). Nevertheless, it is noteworthy that there are two findings that seem inconsistent with the previous studies.

Firstly, self-efficacy is found to negatively relate to beginning teachers' behaviours in terms of providing safe and stimulating learning climates and managing classrooms. But these negative links may weaken and finally turn positive after teachers accumulate certain years of teaching experience. In this case, the finding enriches the previous self-efficacy theories (for a review, see Klassen & Tze, 2014; Zee & Koomen, 2016) by revealing the prevalence of beginning teachers' imprecise perception of their actual capacity in these two domains and by identifying the importance of accumulated experience in lessening such misconception. Comparatively, beginning teachers' evaluation of their actual instructional clarity is relatively more accurate. This may be due to the more tangible indicators (e.g., clear lesson structure, regular checking students' understanding, structured explanation) (Maulana et al., 2020).

Secondly, no evidence was found to uphold the (in)direct relationships between TIOP and differentiated instruction and teaching learning strategies, two behaviour domains that are relatively complex and student-centered. One possible explanation for this could be the measurement instrument used in this study for teachers' self-efficacy, as a higher-order factor, reflecting the general evaluation of their own competence in stimulating and activating teaching, classroom management, and instruction clarity. The lack of domain specificity, particularly in terms of the more complex domains of differentiated instruction and metacognition teaching, may lead to less correspondence between beginning teachers' perception of and actual competence in particular skills. Nevertheless, the empirical validity of the above and additional plausible explanations requires future research.

7 Implication and Limitations

Teaching effectiveness research is not merely concerned with student-centered outcomes. The past decades have witnessed an increasing trend towards paying attention to the significance of teachers in the profession (Keller et al., 2016). Teachers' motivation and well-being as well as the complex mechanisms underlying whether and how they transform such internal qualities into effective teaching behaviour matters. Therefore, this study can serve as a threshold for a fresh view of the inner world of teachers by pointing out a consolidated direction for future research on teachers' psychology-behavior links. Specifically, this empirical study provides some preliminary evidence on the potential benefit TIOP can bring to beginning teachers' well-being and effective teaching behaviour. It is thereby suggested that the theory of TIOP be embedded into the design of initial teacher education (ITE)

and induction arrangements. Nevertheless, findings of the school-level discrepancy between TIOP and self-efficacy that emerge during the recruitment process as well as the teacher-level imprecise perception of actual capacity in certain domains call our attention to a more malleable and differentiated design of such interventions.

During pre-service education, value construction and positive experiencing should be arranged to further nurture student teachers' high meaningfulness and affection for their future career, which is hopefully linked to higher self-efficacy and improved skills in stimulating and activating teaching, classroom management, and clarity of instruction at the individual level. Comparatively, after career entry, schools and mentors are recommended to differentiate their training by providing self-efficacious teachers with TIOP-facilitating intervention (e.g., school visit and enculturation, value construction seminars and workshops) and self-determined teachers with confidence-raising activities (e.g., collaborative lesson planning, peer assessment and communication). It is assumed that such balanced development can not only fashion a more vigorous team of beginning teachers but also advantage their actual teaching behaviour to maximize student learning.

In addition to the school-wide differentiation, teacher education and induction should also offer training that is tailored to teachers' personal characteristics and individual needs (Decker & Rimm-Kaufman, 2008; Joerger & Bremer, 2001). In light of the present research findings, it is suggested that not only teachers' psychological and behaviour profiles (e.g., TIOP, self-efficacy, domain-specific teaching skills) but also personal characteristics (e.g., age, qualification, gender) should be taken into consideration during the design of interventions. Acknowledging the complex interplay of multiple personal factors and how they may influence teachers' well-being and performance in the workplace matters, especially when educators and mentors try to maximize the effectiveness of training and the professional potentials of teachers. In our case, in order to optimize beginning teachers' resiliency to reality shock caused by the discrepancies that emerge among TIOP, well-being, and effective teaching behaviour, additional personalized training and mentoring are recommended.

It is noted that the present study has several limitations. Firstly, this study assessed self-efficacy as a general concept instead of domain-specific self-efficacies (efficacy for instruction, classroom management and student engagement), which to some extent coincides with certain domains of teaching behaviors (e.g., instructional clarity, intensive and activating teaching, classroom management). Therefore, it would be intriguing to further explore the influence of different types of self-efficacy on the related specific domains of teaching behaviour and how such effects mediate the relationships between TIOP and teaching effectiveness. Secondly, the mediation analysis confirmed the assumptions that TIOP constitutes a resource factor and that self-efficacy operates as a mediator between TIOP and basic teaching skills under certain boundary conditions. However, the absence of longitudinal data

makes it impossible to further examine the causality of the relationships. Accordingly, longitudinal or intervention data are needed in future studies to confirm the direction of the effects. Despite the above limitations, the findings support the importance of TIOP for beginning teachers' well-being and effective teaching and demonstrate the moderating effects of teacher-centered background factors. To better understand the complex mechanisms underlying the transformation of TIOP to teaching effectiveness, additional research needs to be conducted. After the hypothesized links are empirically tested in and beyond the current context, the theory-led model constructed in this paper can be validated and applied, as a systematic and generalizable guide, in initial teacher education and teacher induction programs.

Acknowledgement This study was part of the PhD project of the first author, while the second and the third author received a grant from the Dutch Ministry of Education under Grant (OND/ ODB-13/19888). All participants voluntarily participated the project. The first author would like to thank them for their contribution and appreciate Peter Moorer for his support with the management of data.

Appendix

Table 24.A1 English version of three self-reported scales

Sub-scales		Items
TIOP	Professional affection (Adapted from Kunter et al., 2011)	Affection for teaching ET01. I teach [this subject] with great enthusiasm. ET02. I always enjoy teaching students new things. ET03. I enjoy interacting with students. ET04. It's a pleasure to teach.
		Affection for subject ES05. I find my subject exciting and try to convey my enthusiasm to the students. ES06. Engaging in my subject is one of my favorite activities. ES07. I engage in my subject because I enjoy it. ES08. Because engaging in my subject is fun, I wouldn't want to give it up.
	Professional meaningfulness (Adapted from Opdenakker and Maulana 2008)	AM09. ... Because through this work I can achieve my career goals. AM10. ... Because I think it is important for the academic success of my students AM11. ... Because work with interesting challenges gives me satisfaction.

(continued)

Table 24.A1 (continued)

Sub-scales		Items
Perceived self-efficacy (Tschannen-Moran & Hoy, 2001)	Instruction	EIS01. To what extent can you use a variety of assessment strategies? EIS02. To what extent can you provide an alternative explanation or example when students are confused? EIS03. To what extent can you craft good questions for your students? EIS04. How well can you implement alternative strategies in your classroom? EIS05. How well can you respond to difficult questions from your students? EIS06. How much can you do to adjust your lessons to the proper level for individual students? EIS07. To what extent can you gauge student comprehension of what you have taught? EIS08. How well can you provide appropriate challenges for very capable students?
	Classroom management	ECM01. How much can you do to control disruptive behavior in the classroom? ECM02. How much can you do to get children to follow classroom rules? ECM03. How much can you do to calm a student who is disruptive or noisy? ECM04. How well can you establish a classroom management system with each group of students? ECM05. How well can you keep a few problem students from ruining an entire lesson? ECM06. How well can you respond to defiant students? ECM07. To what extent can you make your expectation clear about student behavior? ECM08. How well can you establish routines to keep activities running smoothly?
	Student engagement	ESE01. How much can you do to get students to believe they can do well in schoolwork? ESE02. How much can you do to help your students value learning? ESE03. How much can you do to motivate students who show low interest in schoolwork? ESE04. How much can you assist families in helping their children do well in school? ESE05. How much can you do to help your students think critically? ESE06. How much can you do to foster student creativity? ESE07. How much can you do to get through to the most difficult students? ESE08. How much can you do to improve the understanding of a student who is failing?

Table 24.A2 Fit indices for simple and lower-level mediation models

One-level Model	Robust χ^2 goodness-of-fit						Indirect effect	
	Value	df	CFI	TLI	RMSEA	SRMR	<i>p</i>	95% CI
TIOP on TB1	2.448	4	1.000	1.015	0.000	0.016	.819	[-0.279, 0.337]
TIOP on TB2	10.047*	4	0.977	0.942	0.080	0.038	.724	[-0.385, 0.281]
TIOP on TB3	4.721	4	0.997	0.993	0.027	0.025	.795	[-0.264, 0.382]
TIOP on TB4	2.053	4	1.000	1.019	0.000	0.014	.263	[-0.112, 0.607]
TIOP on TB5	3.696	4	1.000	1.003	0.000	0.019	.204	[-0.107, 0.794]
TIOP on TB6	4.320	4	0.999	0.997	0.018	0.024	.105	[-0.034, 0.865]
TIOP on general TB	2.358	4	1.000	1.016	0.000	0.015	.146	[-0.464, 6.906]
Two-level model	AIC	BIC	Adjusted BIC	Between indirect effect				
				<i>p</i>	95% CI	95% CI (MCMC)		
TIOP on TB1	123.834	217.731	122.742	0.274	[-11.378, 40.090]	[-125.80, 18.24]		
TIOP on TB2	143.664	237.561	142.572	0.971	[-11.583, 11.167]	[-16.07, 7.68]		
TIOP on TB3	147.926	241.823	146.834	0.454	[-3.311, 7.402]	[-4.31, 8.59]		
TIOP on TB4	173.133	267.030	172.041	0.065	[-4.738, 0.139]	[-5.19, -0.23]		
TIOP on TB5	206.568	300.465	205.477	0.696	[-6.693, 10.023]	[-8.43, 8.51]		
TIOP on TB6	226.775	320.672	225.683	0.868	[-27.267, 23.009]	[-6.29, 183.10]		
TIOP on general TB	838.215	932.112	837.123	0.742	[-155.930, 111.059]	[-294.50, 241.30]		

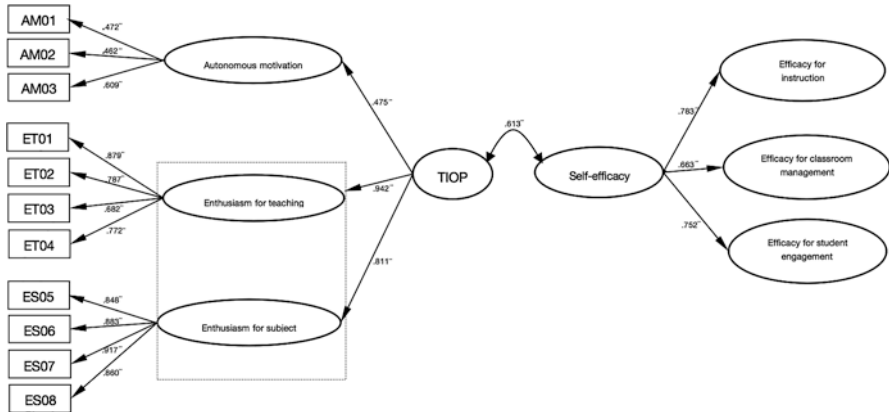


Fig. 24.A1 A model of two focal constructs measured by self-reports for convergent and divergent validity
* $p < .05$

References

- Abbas, Y. (2013). Motivasi intrinsik, motivasi ekstrinsik, kompetensi dan kinerja guru [Intrinsic motivation, extrinsic motivation, competence and teacher performance]. *Humanitas: Indonesian Psychological Journal*, *10*(1), 61–74.
- Ashby, P., Hobson, A. J., Tracey, L., Malderez, A., Tomlinson, P. D., Roper, T., et al. (2008). *Beginner teachers' experiences of initial teacher preparation, induction and early professional development: A review of literature*. DCSF.
- Barr, A. S. (1952). Chapter I: The measurement of teacher characteristics and prediction of teaching efficiency. *Review of Educational Research*, *22*(3), 169–174.
- Bennett, D. A. (2001). How can I deal with missing data in my study? *Australian and New Zealand Journal of Public Health*, *25*(5), 464–469.
- Bernaus, M., Wilson, A., & Gardner, R. C. (2009). Teachers' motivation, classroom strategy use, students' motivation and second language achievement. *Porta Linguarum*, *12*, 25–36.
- Betoret, F. D. (2006). Stressors, self-efficacy, coping resources, and burnout among secondary school teachers in Spain. *Educational Psychology*, *26*(4), 519–539.
- Blatchford, P., Bassett, P., & Brown, P. (2011). Examining the effect of class size on classroom engagement and teacher–pupil interaction: Differences in relation to pupil prior attainment and primary vs. secondary schools. *Learning and Instruction*, *21*(6), 715–730. <https://doi.org/10.1016/j.learninstruc.2011.04.001>
- Blömeke, S., Olsen, R. V., & Suhl, U. (2016). Relation of student achievement to the quality of their teachers and instructional quality. *Teacher Quality, Instructional Quality and Student Outcomes*, 21–50.
- Brigham, F. J., Scruggs, T. E., & Mastropieri, M. A. D. (1992). Teacher enthusiasm in learning disabilities classrooms: Effects on learning and behavior. *Learning Disabilities Research and Practice*, *7*(2), 68–73. Retrieved from <http://www.redi-bw.de/db/ebSCO.php/search.ebSCOhost.com/login.aspx?direct=true&db=eric&AN=EJ443020&site=ehost-live>
- Collie, R. J., Shapka, J. D., Perry, N. E., & Martin, A. J. (2016). Teachers' psychological functioning in the workplace: Exploring the roles of contextual beliefs, need satisfaction, and personal characteristics. *Journal of Educational Psychology*, *108*(6), 788–799.

- de Jesus, S. N., & Lens, W. (2005). An integrated model for the study of teacher motivation. *Applied Psychology: An International Review*, *54*(1), 119–134.
- Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, *19*(2), 109–134.
- Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268.
- Deci, E. L., Olafsen, A. H., & Ryan, R. M. (2017). Self-determination theory in work organizations: The state of a science. *Annual Review of Organizational Psychology and Organizational Behavior*, *4*, 19–43.
- Decker, L. E., & Rimm-Kaufman, S. E. (2008). Personality characteristics and teacher beliefs among pre-service teachers. *Teacher Education Quarterly*, *35*(2), 45–64.
- Dienst Uitvoering Onderwijs. (2015). *Leerlingen per vestiging naar postcode en leerjaar-2014–2015* [Students per location by postcode and year 2014–2015] [Data file]. Retrieved from https://duo.nl/open_onderwijsdata/databestanden/vo/leerlingen/leerlingen-vo-2.jsp
- Dienst Uitvoering Onderwijs. (2016). Alle vestigingen VO [Locations of secondary education schools] [Data file]. Retrieved from https://duo.nl/open_onderwijsdata/databestanden/vo/adressen/
- Dong, Y., & Peng, C. Y. (2013). Principled missing data methods for researchers. *Springerplus*, *2*(1), 222. <https://doi.org/10.1186/2193-1801-2-222>
- Duffin, L. C., French, B. F., & Patrick, H. (2012). The Teachers' sense of efficacy scale: Confirming the factor structure with beginning pre-service teachers. *Teaching and Teacher Education*, *28*(6), 827–834.
- Enders, C. K. (2003). Using the expectation maximization algorithm to estimate coefficient alpha for scales with item-level missing data. *Psychological Methods*, *8*(3), 322–337. <https://doi.org/10.1037/1082-989X.8.3.322>
- Feldon, D. F. (2007). Cognitive load and classroom teaching: The double-edged sword of automaticity. *Educational Psychologist*, *42*(3), 123–137.
- Feng, X., Helms-Lorenz, M., Maulana, R., & Jansen, E. P. (2021). Dutch beginning teachers' intrinsic orientation for the profession: Measurement and consistency during the first year. *Studies in Educational Evaluation*, *70*, 101059.
- Fernet, C., Chanal, J., & Guay, F. (2017). What fuels the fire: Job-or task-specific motivation (or both)? On the hierarchical and multidimensional nature of teacher motivation in relation to job burnout. *Work & Stress*, *31*(2), 145–163.
- Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50.
- Frenzel, A. C., Goetz, T., Stephens, E. J., & Jacob, B. (2009). Antecedents and effects of teachers' emotional experiences: An integrated perspective and empirical test. In P. A. Schutz & M. Zembylas (Eds.), *Advances in teacher emotion research* (pp. 129–151). Springer.
- Gagné, M., Forest, J., Vansteenkiste, M., Crevier-Braud, L., Van den Broeck, A., Aspel, A. K., et al. (2015). The multidimensional work motivation scale: Validation evidence in seven languages and nine countries. *European Journal of Work and Organizational Psychology*, *24*(2), 178–196.
- Gentry, M., Steenbergen-Hu, S., & Choi, B.-Y. (2011). Student-identified exemplary teachers: Insights from talented teachers. *Gifted Child Quarterly*, *55*(2), 111–125. <https://doi.org/10.1177/0016986210397830>
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*, 549–576.
- Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, *10*(3), 229–244.
- Hein, V., Ries, F., Pires, F., Caune, A., Heszteráné Ekler, J., Emeljanovas, A., & Valantiniene, I. (2012). The relationship between teaching styles and motivation to teach among physical education teachers. *Journal of sports science & medicine*, *11*(1), 123–130.

- Helms-Lorenz, M., van de Grift, W., & Maulana, R. (2016). Longitudinal effects of induction on teaching skills and attrition rates of beginning teachers. *School Effectiveness and School Improvement*, 27(2), 178–204.
- Helms-Lorenz, M., Van de Grift, W., Canrinus, E., Maulana, R., & Van Veen, K. (2018). Evaluation of the behavioral and affective outcomes of novice teachers working in professional development schools versus non-professional development schools. *Studies in Educational Evaluation*, 56, 8–20. <https://doi.org/10.1016/j.stueduc.2017.10.006>
- Hong, E., Hartzell, S. A., & Greene, M. T. (2009). Fostering creativity in the classroom: Effects of teachers' epistemological beliefs, motivation, and goal orientation. *The Journal of Creative Behavior*, 43(3), 192–208.
- Hoy, A. W., & Spero, R. B. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and Teacher Education*, 21(4), 343–356.
- Imidayanti, Y., Maulana, R., Helms-Lorenz, M., & Fadhillah, N. (2020). Relationship between teaching motivation and teaching behaviour of secondary education teachers in Indonesia (Relación entre la motivación docente y el comportamiento docente en profesores de educación secundaria en Indonesia). *Journal for the Study of Education and Development*, 43(2), 271–308.
- Jackson, D. L., Teal, C. R., Raines, S. J., Nansel, T. R., Force, R. C., & Burdsal, C. A. (1999). The dimensions of students' perceptions of teaching effectiveness. *Educational and Psychological Measurement*, 59(4), 580–596. <https://doi.org/10.1177/00131649921970035>
- Joerger, R. M., & Bremer, C. D. (2001). *Teacher induction programs: A strategy for improving the professional experience of beginning career and technical education teachers*. NCCTE. http://www.nrccte.org/sites/default/files/publication-files/tchr_indctn_prog.pdf
- Keller, M. M., Hoy, A. W., Goetz, T., & Frenzel, A. C. (2016). Teacher enthusiasm: Reviewing and redefining a complex construct. *Educational Psychology Review*, 28(4), 743–769.
- Kini, T., & Podolsky, A. (2016). *Does teaching experience increase teacher effectiveness? A review of the research*. Learning Policy Institute.
- Klassen, R. M., & Tze, V. M. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review*, 12, 59–76.
- Klusmann, U., Kunter, M., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Teachers' occupational Well-being and quality of instruction: The important role of self-regulatory patterns. *Journal of Educational Psychology*, 100(3), 702–715.
- Kunter, M. (2013). *Cognitive activation in the mathematics classroom and professional competence of teachers results from the COACTIV project*. Springer.
- Kunter, M., & Holzberger, D. (2014). *Loving teaching. Teacher motivation: Theory and practice*, 83–96.
- Kunter, M., Tsai, Y. M., Klusmann, U., Brunner, M., Krauss, S., & Baumert, J. (2008). Students' and mathematics teachers' perceptions of teacher enthusiasm and instruction. *Learning and Instruction*, 18(5), 468–482.
- Kunter, M., Frenzel, A., Nagy, G., Baumert, J., & Pekrun, R. (2011). Teacher enthusiasm: Dimensionality and context specificity. *Contemporary Educational Psychology*, 36(4), 289–301.
- Ladd, H. F., & Sorensen, L. C. (2015). *Do master's degrees matter? Advanced degrees, career paths, and the effectiveness of teachers* (CALDER working paper no. 136). National Center for analysis of longitudinal data in education research.
- Lam, S. F., Cheng, R. W. Y., & Ma, W. Y. (2009). Teacher and student intrinsic motivation in project-based learning. *Instructional Science*, 37(6), 565–578.
- Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404), 1198–1202.
- Lohman, M. C. (2006). Factors influencing teachers' engagement in informal learning activities. *Journal of Workplace Learning*, 18(3), 141–156. <https://doi.org/10.1108/13665620610654577>
- Malmberg, L. E. (2008). Student teachers' achievement goal orientations during teacher studies: Antecedents, correlates and outcomes. *Learning and Instruction*, 18(5), 438–452.

- Marsh, H. W. (1982). The use of path analysis to estimate teacher and course effects in student ratings of instructional effectiveness. *Applied Psychological Measurement*, 6(1), 47–59.
- Marsh, H. W. (1994). Weighting for the right criteria in the instructional development and effectiveness assessment (IDEA) system: Global and specific ratings of teaching effectiveness and their relation to course objectives. *Journal of Educational Psychology*, 86(4), 631–648. <https://doi.org/10.1037/0022-0663.86.4.631>
- Marsh, H. W., & Ware, J. E. (1982). Effects of expressiveness, content coverage, and incentive on multidimensional student rating scales: New interpretations of the Dr. Fox effect. *Journal of Educational Psychology*, 74(1), 126–134. <https://doi.org/10.1037/0022-0663.74.1.126>
- Maulana, R., Helms-Lorenz, M., & Van de Grift, W. (2015). A longitudinal study of induction on the acceleration of growth in teaching quality of beginning teachers through the eyes of their students. *Teaching and Teacher Education*, 51, 225–245. <https://doi.org/10.1016/j.tate.2015.07.003>
- Maulana, R., Helms-Lorenz, M., & Van de Grift, W. (2017). Validating a model of effective teaching behaviour of pre-service teachers. *Teachers and Teaching*, 23(4), 471–493.
- Maulana, R., André, S., Helms-Lorenz, M., Ko, J., Chun, S., Shahzad, A., et al. (2020). Observed teaching behaviour in secondary education across six countries: measurement invariance and indication of cross-national variations. *School Effectiveness and School Improvement*, 1–32.
- Miller, R. T., Murnane, R. J., & Willett, J. B. (2008). Do teacher absences impact student achievement? Longitudinal evidence from one urban school district. *Educational Evaluation and Policy Analysis*, 30(2), 181–200.
- Moè, A., Pazzaglia, F., & Ronconi, L. (2010). When being able is not enough. The combined value of positive affect and self-efficacy for job satisfaction in teaching. *Teaching and Teacher Education*, 26(5), 1145–1153.
- Moulding, N. T. (2010). Intelligent design: Student perceptions of teaching and learning in large social work classes. *Higher Education Research and Development*, 29(2), 151–165. <https://doi.org/10.1080/07294360903470977>
- Natof, T. H., & Romanczyk, R. G. (2009). Teaching students with ASD: Does teacher enthusiasm make a difference? *Behavioral Interventions: Theory & Practice in Residential & Community-Based Clinical Programs*, 24(1), 55–72. <https://doi.org/10.1002/bin.272>
- Opdenakker, M.-C., & Van Damme, J. (2007). Do school context, student composition and school leadership affect school practice and outcomes in secondary education? *British Educational Research Journal*, 33(2), 179–206. <https://doi.org/10.1080/01411920701208233>
- Opdenakker, M.-C., & Maulana, R. (2008). *Overzicht schalen en items leeromgeving docenten project Opdenakker and Maulana Nederland* [Overview of scales and items of teacher learning environments of the Opdenakker and Maulana project]. University of Groningen.
- Opdenakker, M.-C., Maulana, R., & Den Brok, P. (2012). Teacher–student interpersonal relationships and academic motivation within one school year: Developmental changes and linkage. *School Effectiveness and School Improvement*, 23(1), 95–119. <https://doi.org/10.1080/09243453.2011.619198>
- Patrick, B. C., Hisley, J., & Kempler, T. (2000). ‘What’s everybody so excited about?’: The effects of teacher enthusiasm on student intrinsic motivation and vitality. *The Journal of Experimental Education*, 68(3), 217–236. <https://doi.org/10.1080/00220970009600093>
- Pauli, J., Chambel, M. J., Capellari, M. R., & Rissi, V. (2018). Motivation, organisational support and satisfaction with life for private sector teachers in Brazilian higher education. *Higher Education Quarterly*, 72(2), 107–120.
- Pelletier, L. G., Séguin-Lévesque, C., & Legault, L. (2002). Pressure from above and pressure from below as determinants of teachers’ motivation and teaching behaviors. *Journal of Educational Psychology*, 94(1), 186–196.
- Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, 15(3), 209–233.
- Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2011). Alternative methods for assessing mediation in multilevel data: The advantages of multilevel SEM. *Structural Equation Modeling*, 18(2), 161–182.

- Retelsdorf, J., Butler, R., Streblov, L., & Schiefele, U. (2010). Teachers' goal orientations for teaching: Associations with instructional practices, interest in teaching, and burnout. *Learning and Instruction, 20*(1), 30–46.
- Ronfeldt, M. (2012). Where should student teachers learn to teach? Effects of field placement school characteristics on teacher retention and effectiveness. *Educational Evaluation and Policy Analysis, 34*(1), 3–26. <https://doi.org/10.3102/0162373711420865>
- Ross, J. A. (1992). Teacher efficacy and the effects of coaching on student achievement. *Canadian Journal of Education/Revue canadienne de l'éducation, 51*–65.
- Roth, G., Assor, A., Kanat-Maymon, Y., & Kaplan, H. (2007). Autonomous motivation for teaching: How self-determined teaching may lead to self-determined learning. *Journal of Educational Psychology, 99*(4), 761–774.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and Well-being. *American Psychologist, 55*(1), 68–78.
- Ryan, R. M., & Deci, E. L. (2017). Self-determination theory: Basic psychological needs in motivation, development, and wellness. *The Guilford Press*. <https://doi.org/10.1521/978.14625/28806>
- Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology, 99*(3), 611–625.
- Slemp, G. R., Field, J. G., & Cho, A. S. (2020). A meta-analysis of autonomous and controlled forms of teacher motivation. *Journal of Vocational Behavior, 121*, 103459. <https://doi.org/10.1016/j.jvb.2020.103459>
- Sociaal en Cultureel Planbureau. (2014). *Statusscores, peildatum 2014* [Status scores 2014]. Retrieved from https://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores
- Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin, 124*(2), 240–261.
- Tatto, M. T., Peck, R., Schwille, J., Bankov, K., Senk, S. L., Rodriguez, M., ... & Rowley, G. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-MM)*. International Association for the Evaluation of educational achievement. Herengracht 487, Amsterdam, 1017 BT, The Netherlands.
- Taylor, I. M., & Ntoumanis, N. (2007). Teacher motivational strategies and student self-determination in physical education. *Journal of Educational Psychology, 99*(4), 747–760.
- Taylor, I. M., Ntoumanis, N., & Standage, M. (2008). A self-determination theory approach to understanding the antecedents of teachers' motivational strategies in physical education. *Journal of Sport and Exercise Psychology, 30*(1), 75–94.
- Temiz, T., & Topcu, M. S. (2013). Preservice teachers' teacher efficacy beliefs and constructivist-based teaching practice. *European Journal of Psychology of Education, 28*(4), 1435–1452.
- Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*(7), 783–805.
- Tynjälä, P., & Heikkinen, H. L. (2011). Beginning teachers' transition from pre-service education to working life. *Zeitschrift für Erziehungswissenschaft, 14*(1), 11–33.
- Van de Grift, W., Helms-Lorenz, M., & Maulana, R. (2014). Teaching skills of student teachers: Calibration of an evaluation instrument and its value in predicting student academic engagement. *Studies in Educational Evaluation, 43*, 150–159. <https://doi.org/10.1016/j.stueduc.2014.09.003>
- Van den Berghe, L., Soenens, B., Aelterman, N., Cardon, G., Tallir, I. B., & Haerens, L. (2014). Within-person profiles of teachers' motivation to teach: Associations with need satisfaction at work, need-supportive teaching, and burnout. *Psychology of Sport and Exercise, 15*(4), 407–417.
- van der Pers, M., & Helms-Lorenz, M. (2019). Regional school context and teacher characteristics explaining differences in effective teaching behaviour of beginning teachers in The Netherlands. *School Effectiveness and School Improvement, 30*(2), 231–254.

- Van Horn, J. E., Taris, T. W., Schaufeli, W. B., & Schreurs, P. J. (2004). The structure of occupational Well-being: A study among Dutch teachers. *Journal of Occupational and Organizational Psychology*, 77(3), 365–375.
- Van Petegem, K., Aelterman, A., Rosseel, Y., & Creemers, B. (2007). Student perception as moderator for student wellbeing. *Social Indicators Research*, 83(3), 447–463. <https://doi.org/10.1007/s11205-006-9055-5>
- VOION, Arbeidsmarkt en Opleidingsfonds Voortgezet Onderwijs. (2016). *Scenariomodel VO-Basisprognose* [Projections labour market secondary education]. Retrieved from <https://www.scenariomodel-vo.nl/>
- Walberg, H. J., & Paik, S. J. (2000). *Effective educational practices*. International Academy of Education.
- Watt, H. M., & Richardson, P. W. (2008). Motivations, perceptions, and aspirations concerning teaching as a career for different types of beginning teachers. *Learning and Instruction*, 18(5), 408–428.
- Wheless, V. E., Witt, P. L., Maresh, M., Bryand, M. C., & Schrodt, P. (2011). Instructor credibility as a mediator of instructor communication and students' intent to persist in college. *Communication Education*, 60(3), 314–339. <https://doi.org/10.1080/03634523.2011.555917>
- Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, 14(4), 490–495.
- Zee, M., & Koomen, H. M. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher Well-being: A synthesis of 40 years of research. *Review of Educational Research*, 86(4), 981–1015.

Xiangyuan Feng is a PhD candidate at the Department of Teacher Education of the University of Groningen in the Netherlands. Her PhD project focuses on beginning teachers' intrinsic orientation for the profession (TIOP), its influence on teaching skills and academic outcomes, and effective induction interventions to facilitate the development of TIOP. Methodologically, she is interested in structural equation modelling and item response theory.

Michelle Helms-Lorenz is an Associate Professor at the Department of Teacher Education, University of Groningen, The Netherlands. Her research interest covers the cultural specificity versus universality (of behaviour and psychological processes). This interest was fed by the cultural diversity in South Africa, where she was born and raised. Michelle's second passion is education, the bumpy road toward development. Her research interests include teaching skills and well-being of beginning and pre-service teachers and effective interventions to promote their professional growth and retention.

Ridwan Maulana is an associate professor at the Department of Teacher Education, University of Groningen, the Netherlands. His major research interests include teaching and teacher education, factors influencing effective teaching, methods associated with the measurement of teaching, longitudinal research, cross-country comparisons, effects of teaching behaviour on students' motivation and engagement, and teacher professional development. He has been involved in various teacher professional development projects including the Dutch induction programme and school-university-based partnership. He is currently a project leader of an international project on teaching quality involving countries from Europe, Asia, Africa, Australia, and America. He is a European Editor of *Learning Environments Research* journal, a SIG leader of Learning Environments of American Educational Research Association, and chair of the Ethics Commission of the Teacher Education.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 25

The Effects of a Short Self-Access Online Training for Practicum Preparation on the Depths of Reflection of Preservice Teachers



Ye Wang, James Ko , and Peng Wang

Abstract Enhancing preservice teachers' critical reflections on their newly acquired knowledge and experience is crucial for promoting their teaching skills and performance. However, it is a challenging task to teach reflections and increase their depths of reflection. Previous research has succeeded to help preservice teachers reflect in a full-term taught course. However, little empirical research demonstrated the effects of a short self-access online training program on the depths of reflection of preservice teachers. Framed in Ryan and Ryan's (High Educ Res Dev 32(2):244–257, 2013) reflection depth model, this study adopted a quasi-experimental research design to examine the depths of reflection after attending a short online training of four 30-min sessions varied in training session order and session content. Data of 555 reflective statements were identified subsequently in 120 reflective logs of 30 preservice teachers in a teacher education university in northern China. The results showed a significant difference between the experimental and control groups, indicating that a short self-access online training program has beneficial effects on preservice teachers' reflections during practicum preparation. While the depths of the reflective statements identified were relatively shallow, the frequency of the reflective statements did not decrease with their depths. Additionally, topics in the online training sessions significantly affected the depths of preservice teachers' reflections, while the training sequence did not. This study is conducive to designing the relevant online training programmes to promote the depths of reflection of preservice teachers in teacher education programmes.

Y. Wang (✉)

The Education University of Hong Kong, Hong Kong SAR, China

Hiroshima University, Higashihiroshima, Japan

e-mail: s1129913@s.eduhk.hk

J. Ko (✉)

Department of Education Policy and Leadership, The Education University of Hong Kong, Tai Po, Hong Kong

e-mail: jamesko@eduhk.hk

P. Wang

Orthopedics Department, Affiliated Hospital of Hebei University of Chinese Medicine, Shijiazhuang, China

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_25

575

Keywords Reflection · Reflective teaching · Classroom observation · Online training · Preservice teacher

1 Introduction

Building preservice teachers' critical reflections on the understanding and transformation of their acquired knowledge are crucial in shaping their early professional development. However, preservice teachers in general are quite often found to have difficulties and feel disappointed when they taught in the actual classrooms due to the gaps between the acquired knowledge during teacher education and practicum experience (Korthagen et al., 2006). Defined as "deliberate thinking about action with a view to its improvement" (Hatton & Smith, 1995, p. 40), reflection is a valued emphasis in the current field experience requirements of teacher education programmes because preservice teachers are expected to develop the "ability to facilitate learning and talk meaningfully about their practice" (Tyrrell et al., 2013, p. 15). Peer dialogues in a virtual learning environment of preservice teachers from four countries on practical teaching issues indicated different reflected depths of teacher education and practicum (Wang et al., 2020). Classroom observation, which provides a direct way to observe and evaluate teachers' teaching behaviours, is a vital tool for teacher evaluation and professional development (Martinez et al., 2016).

Therefore, reflective teaching and classroom observation can be two practical approaches for preservice teachers to improve reflection depth and ultimately achieve quality teaching. This study describes a pilot study examining the impact of a short self-access online training in reflective teaching (RT) and classroom observation (CO) on preservice teachers' depths of reflection before they starts their practicum practice.

2 Literature Review

2.1 *Reflective Teaching*

Reflective teaching refers to teachers' reflections of their teaching practice in the classroom, especially the problems they meet during teaching, to put forward appropriate strategies and methods to resolve them (Schön, 1983). However, problem-solving is not the only feature of reflective teaching. The knowledge and experience that preservice teachers have acquired in the past may not provide sufficient support for the current teaching situation. Preservice teachers need to transform from relying on the prior knowledge and experience to actively achieving new knowledge and creating new thoughts. Reflection allows preservice teachers to adopt the newly reconstructed knowledge into practice, thus further enhancing their teaching skills and promoting their practice. For instance, Lee (2005) believed that through teaching reflection, preservice teachers could continuously enrich their teaching

Table 25.1 Model for teaching and assessing reflective learning (Ryan & Ryan, 2013)

Reflection levels	Questions to get started
Reporting and Responding	Report what happened or what the issue or incident involved. Respond to the incident or issue by making observations, expressing one's opinion, or asking questions.
Relating	Relate or make a connection between the incident or issue and one's own skills, professional experience, or discipline knowledge.
Reasoning	Highlight in detail significant factors underlying the incident or issue. Explain and show why they are essential to understand the incident or issue. Refer to relevant theory and literature to support one's reasoning.
Reconstructing	Reframe or reconstruct future practice or professional understanding.

knowledge, develop their teaching competence, apply the constructed new knowledge and accumulate experience in teaching practice.

However, preservice teachers need to consider a broader range of teaching to accomplish teaching effectiveness through an in-depth reflection. Moon (2007), for instance, classified four reflective writing levels: descriptive writing, descriptive writing with some reflection, descriptive reflective writing, and in-depth reflective writing. Considering how and the depth that learners reflect on their teaching practice, Ryan and Ryan (2013) created a Model for Teaching and Assessing Reflective Learning (TARL) for students and teachers to develop their critical thinking levels of reflection in tertiary education. TARL involves four hierarchical levels of reflection: reporting and responding, relating, reasoning, and reconstructing (see Table 25.1). TARL provides a holistic understanding of the gradual progress of reflections from elementary to profound by preservice and in-service teachers. The TARL model is beneficial for teachers to improve their teaching performance by describing and responding to a simple question using related theories to explain and better resolve the issues (Barton & Ryan, 2014).

2.2 Classroom Observation

Classroom observations also contribute in generating deeper reflection on teaching performance. With an accurate teaching and learning situation, preservice teachers could objectively observe what happens in the classroom through classroom observation (CO). Peer observation, which is also beneficial to teachers' professional development (O'Connell et al., 2000), helps preservice teachers reflect their teaching performance and form new insights.

Different classroom observation instruments have been used to evaluate teachers' teaching practices. As a widely-researched instrument, the International Comparative Analysis of Learning and Teaching (ICALT) aims to study teaching behaviours and examine teaching behaviour growth (Van de Grift, 2007). The ICALT instrument has been adopted to help preservice teachers to improve their teaching practice during the teacher education period (Maulana & Helms-Lorenz, 2016; Maulana et al., 2017). Research shows that the ICALT stage model provides

an appropriate description of the development of effective teaching for most teachers and each teacher's current teaching skills (Van der Lans et al., 2017). Additionally, it has been verified that the ICALT instrument is invariance for measuring effective teaching across five different countries (Maulana et al., 2019).

2.3 Online Training of Preservice Teachers

With the popularity of massive open online courses (MOOCs), universities are eager to supplement existing curricula and self-regulating learning with online modules. With the rapid development of high technology, online learning and online training have contributed to teachers' teaching reflection and professional development (Bates et al., 2016). In-service teachers considered their teaching benefitted from a one-year teacher online training program (Krammer et al., 2006).

Moreover, online learning through high technology, such as mobile phones and other wireless technologies, offers a fragmented time and relaxed atmosphere for encouraging preservice teachers to engage in the learning activities (Becker et al., 2018). An online learning platform allows preservice teachers to learn asynchronously with more autonomy and selectivity without time and space limit.

2.4 Practicum Preparation in the Chinese Context

Generally, both primary and secondary school teaching qualifications take four years' study. To cultivate research-oriented teachers with a solid basis of theoretical knowledge and teaching practice, a few top teacher education universities offer two to three years of graduate study by selecting some excellent students.

Teaching reflection has been emphasised in teacher education programmes. According to the new curriculum of teacher education program (Ministry of Education [MOE] of China, 2011), preservice teachers need to have the ability to critically think about their learning and teaching, thus becoming reflective practitioners. They also need to prepare themselves as life-long learners, thus to continuously promoting their knowledge and teaching skills through the whole teaching career. During practicum preparation, they are required to deepen their understanding of the specific subject knowledge as well as pedagogical knowledge and develop the ability to launch reflective teaching and solve teaching problems through formal courses (i.e. teaching case study, classroom observation of high-quality class and famous teachers), and various learning activities (i.e. learning community, group discussion). MOE of China (2014) proposed setting up a new trinity mode in which teacher education universities, local governments and local schools cooperate to strengthen teacher preparation of preservice teachers. Preservice teachers can be well prepared during practicum practice.

There is a lack of research that considers the effects of an online training programme on the depths of preservice teachers' reflections. Therefore, this study aimed at exploring the effects of a short self-access online training in reflective teaching and classroom observation on preservice teachers' depths of reflection. The research questions addressed are as follows:

1. Can online training in reflective teaching and classroom observation enhance preservice teachers' depths of reflection?
2. Do reflection topics make a difference in preservice teachers' depths of reflection?
3. Does the order of the training sequence make a difference in preservice teachers' depths of reflection?

3 Methodology and Research Design

This study adopted a quasi-experimental research design with two different methods to examine the effects of a short self-access online training on preservice teachers' depth of reflection.

3.1 Participants

Thirty preservice teachers were recruited from a teacher education university in Hebei Province in northern China. All of them were in the second semester of their junior year during the data collection. Their ages varied from 18 to 24 years old. Three were male and twenty-seven were female. Their majors were classified into four majors: math and science studies, language studies (Chinese and English language study), primary education study, and others (i.e. History, Geography, Physical Education). In their future practicum practice, seven participants would be assigned to primary schools, and twenty-three to secondary schools based on their majors. However, their acquired knowledge was similar during practicum preparation. The participants were randomly divided into three groups: the control group without any training and two experimental groups, the Reflective Teaching-Classroom Observation Group (RT-CO Group) and the Classroom Observation-Reflective Teaching Group (CO-RT Group), differed in the sequence of the two training sessions (i.e., RT and CO). Each group has ten participants. Table 25.2 shows the details of the participants in each group. All participants joined voluntarily and were briefed on the research aim and procedures before submitting their consent forms. All information related to the participants was treated anonymously and confidentially.

Table 25.2 Information of participants

Group	Major	School level for practicum preparation	Number	Total
RT-CO Group	Math & Science	Secondary	3	10
	Language Studies	Secondary	4	
	Primary Education	Primary	1	
	Other Majors	Secondary	2	
CO-RT Group	Math & Science	Secondary	4	10
	Language Studies	Secondary	2	
	Primary Education	Primary	4	
	Other Majors	Secondary	0	
Control Group	Math & Science	Secondary	5	10
	Language Studies	Secondary	2	
	Primary Education	Primary	2	
	Other Majors	Secondary	1	
Total				30

3.2 Instruments

3.2.1 Online Training

The content of the training session was designed based on the relevant literature of RT and CO. For instance, the RT sessions were based on the studies on reflection and reflective teaching (e.g., Moon, 2007; Hall & Simeral, 2015) and collaborative reflection (e.g., Prilla & Renner, 2014; Wang & Quek, 2015); The CO sessions were based on the studies on classroom observation, effective teaching, and inspiring teaching (e.g., Borich, 2010; Van de Grift, 2014; Sammons et al., 2014, 2016; Ko et al., 2019).

Four narrated PowerPoints were developed on two themes, two on RT and two on CO. “What do preservice teachers need to know about RT?”, “How can you become a reflective teacher?”, “What do preservice teachers need to know about CO?”, “How to do classroom observation?”. The training sessions provided various learning activities to motivate preservice teachers to learn autonomously. The PowerPoints were designed initially in English and then translated into Chinese to make them more accessible for the participants. Figures 25.1 and 25.2 show some screenshots of the PowerPoints.

3.2.2 Topics as Stimulation for Reflections

We explored the reflection task effects on reflection because preservice teachers understand and reflect on different teaching contexts using scenarios during teacher education (Snoek, 2003; Aubusson & Schuck, 2013). Thus, after each training session, participants were given two topics to stimulate their reflections to write a log for each. In the RT training sessions, participants were asked to comment on a math

Key Behaviors Essential for Effective Teaching:

- 1) Lesson clarity
- 2) Instructional variety
- 3) Teacher task orientation
- 4) Engagement in the learning process
- 5) Students success rate

What do **EFFECTIVE** teachers believe? What attitudes do they hold about **EFFECTIVE** teaching and learning?

How do you know all this? Which of these beliefs and attitudes are **LEARNABLE**? Which of these beliefs and attitudes are **OBSERVABLE**?

Effective Teachers

(Borich, 2010; Scheuerman, 2013)

有效教学的关键行为

1. 课程清晰度
2. 教学多样化
3. 教师任务取向
4. 学生参与度
5. 学生成功率

教师的教学信念是什么? 他们对有效教学的信念是什么? 你如何知道这些的? 哪些信念和态度是更容易掌握的? 哪些信念和态度是可以观察到的?

Effective Teachers

(Borich, 2010; Scheuerman, 2013)

Model for Teaching and Assessing Reflective Learning (TARL) (Ryan & Ryan, 2013)

反思性学习的教学与评价模式 (TARL) (Ryan & Ryan, 2013)

Fig. 25.1 The screenshots of slides of online training in reflective teaching (English version VS Chinese version)

Applications of classroom observation (Ko, 2010)

课堂观察的应用 (Ko, 2010)

Implicit matters in your lesson observation

Do you agree with the following statement?

1) Write your own words can give you insights on your teaching.

2) Lesson observation shows your teaching performance only.

- It may show your content-specific knowledge.
- It may show your teacher-student relationship.
- It may show how well your lesson planning is.

课堂观察中隐含的事物

1) 课堂观察能观察到你的教学表现。

- 可以观察到你对特定内容知识的掌握
- 可以观察到你与学生的关系
- 可以观察到你的教学计划做得怎么样

Fig. 25.2 The screenshots of slides of online training in classroom observation (English version VS Chinese version)

teacher’s teaching reflection (Topic 1, Fig. 25.3) and write a reflective log on their own limited teaching experience (Topic 2, Fig. 25.4). In CO training sessions, participants were asked to write whatever they wanted to discuss after observing a teacher teaching insects (Topic 3, Fig. 25.5) and two overseas teachers teaching Geography and Math (Topic 4, Fig. 25.6).

Teaching reflection case

Alisa is a math teacher in the third grade of elementary school. Before class on Monday, she thought about what she would teach today, and then decided to do as usual during class. She first explained the concept of decimal to the students, and then let the students complete the exercises in the textbook. But most of the students behaved very inactively and made many mistakes. The math class that day was not going well. Teacher Alisa knows that if she spends more time preparing lessons and more interactive sessions during the lecture, she may be more successful. That afternoon, Teacher Alisa also gave feedback to the head of the teaching and research team: "The students in my class lack basic concepts, and some people seem to have never heard of one hundred. I think this unit will be very difficult to perform. When it comes to the unit tests, they won't get a good score." (Hall & Simeral, 2015)

Please comment on the teacher's reflection. You can consider (but not limited to) the following issues:
 What level of reflection is this teacher?
 What did she reflect on?
 What is the problem she is facing?
 Why do these problems occur?
 If you are this teacher, how do you solve this problem? Please list at least 3 solutions and briefly explain.

◆ Word limit: at least 100 words

Fig. 25.3 Topic 1- A math teacher's teaching reflection

Imagine that you are a teacher. Please write a self-reflection diary based on the following questions (but not limited to):

1. Why did you enter the teacher industry? As a teacher, what are your advantages?
2. In teaching, what problems do you think you will face? Please give an example.
3. In your opinion, why does this kind of problem occur?
4. How do you solve such problems?
5. In your class, those students with good grades, what are the reasons for their success?
6. Which type of students in your class is more difficult to study? What are their special needs?
7. What adjustments can you make to meet these special needs?
8. How do you judge the teaching effect of a class? Why?

Writing examples:
 I think it is difficult for me to complete all the teaching tasks of a class, because I will usually spend a long time to explain with examples. I think I will let students spend most of my time doing exercises, but the class atmosphere may be dull! I want to raise the enthusiasm of students by asking questions...

Word limit: at least 100 words

Fig. 25.4 Topic 2- A self-reflection diary

3.3 Training Session Sequence

We could not find any literature on the effects of learning RT and CO in different sequences. Both topics were not formally taught in the university of the participants. All participants of the two experimental groups were asked to go through a training session of two PowerPoints in two weeks. However, the training sequence was different for each group (see Table 25.3). The RT-CO Group took two sessions of RT first and then two sessions of CO; the CO-RT Group took the training sessions in reverse order. The participants in the two experimental groups wrote the reflective logs according to the training sequence.

Just imagine, you've just observed a lesson about "Insects". Below is the beginning part of this lesson. After lesson observation, what do you want to discuss with this teacher?

Teacher: I'm going to give you the little word 'insect'. Immediately in your mind there's a picture of something, I expect. There is in mine. What sort of picture have you got Cassandra?
Pupil 1: A spider.
Teacher: OK, you think of a spider. You keep the spider there. Catherine, what about you?
Pupil 2: (no response.)
Teacher: When I say 'insect' what do you immediately think of – an insect?
Pupil 3: A ladybird.
Teacher: Yes, that's right.
Pupil 4: A worm.

Teacher: Yes – anything else?
Pupil 3: A snail.
Teacher: How do insects move around, Peter?
Pupil 2: Legs.
Teacher: How many legs has an insect got?
Pupil 2: Six.
Teacher: Yes, six, but do insects get around any other way?
Pupil 2: Some insects fly.
Teacher: Yes, some insects use wings. Can you think of an insect that flies?
Pupil 2: An eagle.
Teacher: An eagle? Is that an insect? No, it's a bird. A bird is definitely not an insect.

Word limit: at least 100 words

Wragg, T. (2011)

41

Fig. 25.5 Topic 3- A teaching case of insects

Please observe lesson videos of two overseas teachers:

1. Can you tell the differences in teaching among the Chinese teachers and the two teachers in the videos?
2. What advantages and disadvantages do they have?
3. Which teaching strategies/methods/styles do you think can be used in your own teaching in the future?



Word limit: at least 100 words

36

Fig. 25.6 Topic 4- Classroom observation

Table 25.3 Training sequence for three groups

Group	Online training sequence	
	First two sessions	Last two sessions
RT-CO Group	RT (Topic 1, 2)	CO (Topic 3, 4)
CO-RT Group	CO (Topic 3, 4)	RT (Topic 1, 2)
Control Group	No training (Topic 1, 2) (Topic 3, 4)	

Note: the sequence to write the reflective logs for each group shows in the brackets

Training sessions with reflective logs were delivered online to participants via WeChat, a Chinese instant messaging system. Each time, all participants were asked to finish learning an online training session before submitting a reflective log in three days. Although the control group did not take the online training, they still needed to write reflective logs like their peers in the two experimental groups. The order of writing reflective logs for the participants in the Control Group was the same as the RT-CO Group.

3.4 Data Analysis

First, an in-depth qualitative dialogue analysis (Hennessy et al., 2016) was conducted to determine the depths of reflection in the reflective logs. The four hierarchical levels of TARL (Ryan & Ryan, 2013) were adopted to categorize the depths of every reflective statement found. Table 25.4 shows the coding descriptions with

Table 25.4 Code descriptions of preservice teachers' reflective logs

Category	Code	Label	Description and example excerpts of reflective statements from reflective logs
Reporting and Responding	RL1	Reporting	Report what happened or what the issue or incident involved. Respond to the incident or issue by making observations, expressing an opinion, or asking questions. For example, <i>"The students did not understand the decimal system the teacher taught in class, so they made lots of mistakes in the exercises."</i>
Relating	RL2	Draw relationships with existing knowledge or experience	Relate or make a connection between the incident or issue and preservice teachers' own skills, professional experience, or discipline knowledge. Refer to the viewpoints of peers/ colleagues/ experts. For example, <i>"From my knowledge about student engagement, the teacher should motive students' interests with the aids of pictures/ videos."</i>
Reasoning	RL3	Explore or explain the factor of the phenomenon	Highlight in detail significant factors underlying the incident or issue. Explain and show why the influential factors are essential to an understanding of the incident or issue. For example, <i>"I think the teacher was weak in classroom management and instruction clarity and student engagement because she did not sufficiently prepare the lesson."</i>
Reconstructing	RL4	Reframe experience with a theoretical perspective for future pedagogy	Based on the theories by someone, the preservice teacher should provide opportunities for a personalised learning experience to increase or student autonomy. For instance, <i>"After discussing with my mentor talking to student engagement, I will stimulate students' engagement through positive interactions with them in the future."</i>

some examples of excerpts from reflective logs. The first author split the reflective logs into reflective statements line by line and then coded and categorized them according to the code descriptions and examples. Another coder from the same project team verified the splitting of reflective statements.

The second coder coded 10% of the total materials. The interrater reliability was high using Krippendorff's (1980) Alpha ($\alpha = 0.88$). The first author made the final decision of the coding disagreements and coded the remaining reflective statements.

Second, chi-square tests were conducted with SPSS 25 to determine how the online training might affect the generation of reflective statements in different depths and whether various topics and training sequences might matter.

4 Findings

In total, 555 reflective statements were identified from 120 reflective logs of 30 participants. The descriptive statistics in Table 25.5 showed that the participants in the two experimental groups generated more reflective statements ($N = 231$, $N = 180$, respectively) than that of the Control Group ($N = 144$). The total mean score of preservice teachers' depths of reflection was 1.29 ($SD = 0.68$). The depths of reflection of two experimental groups ($M = 1.41$, $SD = 0.79$; $M = 1.26$, $SD = 0.66$; respectively) were slightly higher than that of the Control Group ($M = 1.13$, $SD = 0.45$).

4.1 Depths of Reflection of Preservice Teachers among Groups

In general, the reflective statements tend to be at the *Reporting and Responding* level ($N = 466$, 84%), rather than the *Relating* level ($N = 17$, 3.1%) and the *Reasoning* level ($N = 72$, 13.0%). No *Reconstructing* reflective statements were found. Table 25.6 shows that the percentages of *Reporting and Responding* statements (the RT-CO Group: 38.6%, the CO-RT Group: 33.3%, the Control Group: 28.1%), and *Reasoning* statements (the RT-CO Group: 61.1%, the CO-RT Group: 30.6%, the Control Group: 8.3%) in two experimental groups were higher than that of the Control Group. However, the percentage showed in *Relating* statements was at the same level between the RT-CO Group (41.2%) and the Control Group (41.2%),

Table 25.5 Descriptive statistics of the depths of reflection

Group	Number of reflective statements	M	SD
RT-CO Group	231	1.41	.79
CO-RT Group	180	1.26	.66
Control Group	144	1.13	.45
Total	555	1.29	.68

Table 25.6 Cross-tabulation of depths of reflection among groups

Depths of reflection	RT-CO Group		CO-RT Group		Control Group		Subtotal	
	Count	%	Count	%	Count	%	Count	%
Reporting & responding	180	77.9	155	86.1	131	91.0	466	84.0
Draw relationships with existing knowledge or experience	7	3.0	3	1.7	7	4.9	17	3.1
Explore or explain the factor of the phenomenon	44	19.0	22	12.2	6	4.2	72	13.0
Total	231	100	180	100	144	100	555	100

Note: $\chi^2(4, N = 555) = 19.87, p = .001$

One cell (11.1%) have an expected count of less than 5. The minimum expected count is 4.41

Table 25.7 Distribution of reflective statements at different depths of reflections by different topics

Depths of reflection	Topic 1		Topic 2		Topic 3		Topic 4		Subtotal	
	Count	%	Count	%	Count	%	Count	%	Count	%
Reporting & responding	135	82.3	118	88.1	71	79.8	142	94.0	466	86.6
Explore or explain the factor of the phenomenon	29	17.7	16	11.9	18	20.2	9	6.0	72	13.4
Total	164	100	134	100	89	100	151	100	538	100

Note: $\chi^2(3, N = 538) = 13.63, p = .003$

while the CO-RT Group showed low percentage (17.6%). The percentages of participants' reflection depths was significantly different by group, $\chi^2(4, N = 555) = 19.87, p = 0.00$.

Based on the adjusted Z scores, a post hoc test showed that only the RT-CO Group was significantly different from the Control Group in *Reasoning*, $p < 0.00$. Moreover, a significant difference was shown between the appropriate proportions of the *Reporting and Responding* statements and *Reasoning* statements in the RT-CO Group, $p < 0.00$. The participants in the RT-CO Group were more likely to generate reflective statements related to the *Reasoning* statements.

4.2 Comparison of the Depths of Preservice Teachers' Reflection by Different Topics

Table 25.7 shows that the participants generated more reflective statements in Topic 1, Topic 2 and Topic 4 ($N = 164, 30.5\%$; $N = 134, 24.9\%$; $N = 151, 28.1\%$; respectively), while fewer in Topic 3 ($N = 89, 16.5\%$). The proportion of the *Reporting and Responding* reflective statements ($N = 466, 86.6\%$) was the most prevalent in each topic, whereas the proportions of statements in the *Reasoning* category were relatively small ($N = 72, 13.4\%$). No statements in Topic 3 and Topic 4 were found in the *Relating* category. The depths of reflection of preservice teachers were significantly different by the topics, $\chi^2(3, N = 538) = 13.63, p = 0.00$.

Based on the adjusted Z scores, a post hoc test demonstrated that the proportions of reflection depths in Topic 4 differed significantly, $p < 0.00$. A relatively more significant proportion of the *Reporting and Responding* statements was shown in Topic 4.

4.3 Comparison of Depths of Reflection Between the Two Experimental Groups

Table 25.8 shows that the total count of reflective statements of the two topics on RT in the RT-CO Group ($N = 65$, $N = 54$, respectively) was higher than the CO-RT Group ($N = 22$, $N = 49$, respectively) whose topics were in CO after the participants finished the first two online training sessions. The total count of reflective statements of the two topics on RT in the CO-RT Group ($N = 64$, $N = 45$, respectively) was almost the same as the RT-CO Group ($N = 47$, $N = 65$, respectively) whose topics were in the CO. According to the results showed within the RT-CO Group, the percentages of reflective statements for two themes were similar (RT: 51.5%, CO: 48.4%). According to the results showed within the CO-RT Group, the frequency of reflective statements for the CO theme was 39.4%. However, the percentage of reflective statements was increased to 60.6% after finishing the last two online training sessions in RT. However, there was no significant difference between these two experimental groups by different topics, $\chi^2 (3, N = 411) = 5.89, p = 0.12$.

5 Conclusion and Discussion

This study explored the impact of online training in reflective teaching and classroom observation on the depths of reflection of preservice teachers. The results have generally verified the beneficial effects of the short self-access online training program and different topics, a no significant association with the training sequence, and a lack of depth in reflections despite online training.

Table 25.8 Distribution of reflective statements in different topics between two experimental groups

Variable	RT-CO Group		CO-RT Group		Subtotal	
	Count	%	Count	%	Count	%
Topic 1 (RT1)	65	15.8	64	15.6	129	31.4
Topic 2 (RT2)	54	13.1	45	10.9	99	24.1
Topic 3 (CO1)	47	11.4	22	5.4	69	16.8
Topic 4 (CO2)	65	15.8	49	11.9	114	27.7
Total	231	100	180	100	411	100

Note: $\chi^2 (3, N = 411) = 5.89, p = .12$

RT Reflective Teaching, CO Classroom Observation

5.1 Lack of Depth in Reflection in Chinese Preservice Teachers

The results showed that most of the participants' reflective statements were at the *Reporting and Responding* level, indicating that all participants' depths of reflection were relatively shallow. Despite the sample of excellent preservice teachers, we may expect the issue to be more worrisome in average and lower quality of teacher education based on student in-take. The nature of the self-access online training might explain the low performance because the preservice teachers were doing it without credit, and the online training sessions might hamper motivation to provide reflections the best they could. External factors in the social context could affect intrinsic motivation that stimulates people to produce satisfactory results (Dörnyei & Ushioda, 2013).

Contrary to the prediction of the TARL model (Ryan & Ryan, 2013), the Reasoning level showed a higher frequency than the Relating level. The preservice teachers in the Chinese context may have some difficulties in developing reflections. The Chinese preservice teachers tended to be more aware of pointing out the main elements of teaching problems. Still, they could not link the incidents that happened in the classroom with their theoretical knowledge. Teacher education reform has been promoted and deepened in China, but there are still some problems. As preservice teachers seldom have opportunities to teach in an authentic classroom, it is difficult to integrate their learned knowledge with practical practice during the initial teacher education stage. Their teaching reflection should also be improved (Chen, 2008; Li & Qin, 2015). It suggested that preservice teachers should be encouraged to critically think about their learning and teaching during the teacher education programme. Thus they could achieve higher teaching quality with developed reflections. Reflective skills significantly impact students' perception of integration theory with practice (Hatlevik, 2012).

5.2 Beneficial Effects of the Online Practicum Preparation and Tasks of Instructional Design

The results showed a significant difference between the experimental and control groups, suggesting that the online training in reflective teaching and classroom observation could enhance preservice teachers' reflections. The knowledge the preservice teachers acquired from such online training was beneficial for improving their reflections. It suggested that a short self-access online training program could positively support preservice teachers' reflection during teacher education. Maulana and his colleagues (2015) have demonstrated that novice teachers' teaching skills could be remarkably improved if they received support from teacher induction programmes, such as formal and informal teacher training and mentors' guidance. Caywood and Duckett (2003) have found out that there were no significant

differences between online teaching and on-campus teaching in student teachers' learning outcome in teacher education.

Our results showed that the topics provided for preservice teachers after the online training sessions differed significantly, indicating different tasks may also affect the depths of preservice teachers' reflections. Different tasks and scenarios could stimulate preservice teachers to think about the actual teaching situations and consider how to teach in the real classroom during practicum preparation. Student teachers got higher scores practising the given tasks, and their pedagogical knowledge improved before experiencing the actual classroom (Badiee & Kaufman, 2014). The internship experience of preservice teachers could be enriched through proper reflection tasks (Oner & Adadan, 2011). The results indicated that observing lesson videos could trigger preservice teachers to reflect more during their teacher training stage. Preservice teachers could have more profound reflections via an online video-case study in a teacher training program (Bayram, 2012).

5.3 Primacy of Reflection Training

The results showed that there was no significant difference between the RT-CO Group and the CO-RT Group by different topics. Nevertheless, the CO-RT Group caught up with the RT-CO Group after finishing the last two online training sessions in RT. Additionally, these two experimental groups generated more reflective statements after finishing the online training sessions in RT than CO. This result indicated that the training sequence might make a difference. The training sessions in reflective teaching is conducive for preservice teachers to develop their reflective ability through knowledge construction. The development of preservice teachers' thinking towards their teaching practice and the acquired knowledge during initial teacher education can improve their teaching quality effectively. It has been demonstrated that reflection plays a vital role in initial teacher education (Pedro, 2005; Lee, 2008; Williams & Grudnoff, 2011). Teachers could achieve teaching effectiveness through integrating their enhanced understanding in teaching with better actions by reflection, and they could regard it as the foundation of the subsequent reflection (Ash & Clayton, 2004).

5.4 Limitations

We also acknowledge the limitations of this study. This study has been verified for the effects of our short self-access online training sessions. Future studies could explore whether preservice teachers' reflection depth could be improved if they take the short online training as a part of credited courses. In this study, preservice teachers' reflection levels were relatively shallow. Future studies could adopt collaborative reflection, such as group discussion, to stimulate their reflection depths.

Additionally, due to the pandemic of COVID-19, the schedule for online training became very tight. The online training was conducted within two weeks to finish all training sessions before preservice teachers started their teaching practicum. Therefore, they may not have enough time to reflect and consolidate what they have learned during online training. Future research could extend the length of online training for preservice teachers to have sufficient time to develop reflective skills.

5.5 *Significance and Implications for Teacher Educators and Instruction Designers*

By exploring the impact of short self-access online training sessions designed to stimulate reflections, this study has contributed to a fresh understanding of their strengths and limitations. This study contributes to the instructional design of reflection training with different tasks and their potentials in a teacher education programme. Moreover, this study is also conducive to encourage preservice teachers to reflect more and deeper on their teaching practice and ultimately develop professionalism based on solid reflective practices.

References

- Aubusson, P., & Schuck, S. (2013). Teacher education futures: Today's trends, tomorrow's expectations. *Teacher Development*, 17(3), 322–333.
- Ash, S. L., & Clayton, P. H. (2004). The articulated learning: An approach to guided reflection and assessment. *Innovative Higher Education*, 29(2), 137–154.
- Badiee, F., & Kaufman, D. (2014). Effectiveness of an online simulation for teacher education. *Journal of Technology and Teacher Education*, 22(2), 167–186.
- Barton, G., & Ryan, M. (2014). Multimodal approaches to reflective teaching and assessment in higher education. *Higher Education Research & Development*, 33(3), 409–424.
- Bates, M. S., Phalen, L., & Moran, C. G. (2016). If you build it, will they reflect? Examining teachers' use of an online video-based learning website. *Teaching and Teacher Education*, 58, 17–27.
- Bayram, L. (2012). Use of online video cases in teacher training. *Procedia-Social and Behavioral Sciences*, 47, 1007–1011.
- Becker, S. A., Brown, M., Dahlstrom, E., Davis, A., DePaul, K., Diaz, V., & Pomerantz, J. (2018). *NMC horizon report: 2018 higher education edition*. Educause.
- Borich, G. D. (2010). *Effective teaching methods: Research-based practice*. Pearson Prentice Hall.
- Caywood, K., & Duckett, J. (2003). Online vs. on-campus learning in teacher education. *Teacher Education and Special Education*, 26(2), 98–105.
- Chen, X. (2008). The role of theories in teachers' professional development. *Peking University Education Review*, 6(1), 39–50.
- Dörnyei, Z., & Ushioda, E. (2013). *Teaching and researching: Motivation*. Routledge.
- Hall, P., & Simeral, A. (2015). *Teach, reflect, learn: Building your capacity for success in the classroom*. ASCD.

- Hatlevik, I. K. R. (2012). The theory-practice relationship: Reflective skills and theoretical knowledge as key factors in bridging the gap between theory and practice in initial nursing education. *Journal of Advanced Nursing*, 68(4), 868–877.
- Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, 11(1), 33–49.
- Hennessy, S., Rojas-Drummond, S., Higham, R., Márquez, A. M., Maine, F., Ríos, R. M., et al. (2016). Developing a coding scheme for analysing classroom dialogue across educational contexts. *Learning, Culture and Social Interaction*, 9, 16–44.
- Ko, J., Sammons, P., Maulana, R., Li, V., & Kyriakides, L. (2019, April 5–9). *Identifying inspiring versus effective teaching: How do they link and differ?* [Paper presentation]. American Educational Research Association Annual Meeting, Toronto, ON, Canada.
- Korthagen, F., Loughran, J., & Russell, T. (2006). Developing fundamental principles for teacher education programs and practices. *Teaching and Teacher Education*, 22(8), 1020–1041.
- Krammer, K., Ratzka, N., Klieme, E., Lipowsky, F., Pauli, C., & Reusser, K. (2006). Learning with classroom videos: Conception and first results of an online teacher-training program. *ZDM*, 38(5), 422–432.
- Krippendorff, K. (1980). *Content analysis. An Introduction to its methodology* (The Sage Comtext Series). Sage.
- Lee, H. J. (2005). Understanding and assessing preservice teachers' reflective thinking. *Teaching and Teacher Education*, 21(6), 699–715.
- Lee, I. (2008). Fostering preservice reflection through response journals. *Teacher Education Quarterly*, 35(1), 117–139.
- Li, X. H., & Qin, J. (2015). The development and improvement of intern students' PCK in apprenticeship. *Educational Research*, 12, 141–145.
- Martinez, F., Taut, S., & Schaaf, K. (2016). Classroom observation for evaluating and improving teaching: An international perspective. *Studies in Educational Evaluation*, 49, 15–29.
- Maulana, R., & Helms-Lorenz, M. (2016). Observations and student perceptions of the quality of preservice teachers' teaching behaviour: Construct representation and predictive quality. *Learning Environments Research*, 19(3), 335–357.
- Maulana, R., Helms-Lorenz, M., & Van de Grift, W. (2015). A longitudinal study of induction on the acceleration of growth in teaching quality of beginning teachers through the eyes of their students. *Teaching and Teacher Education*, 51, 225–245.
- Maulana, R., Helms-Lorenz, M., & Van de Grift, W. (2017). Validating a model of effective teaching behaviour of preservice teachers. *Teachers and Teaching*, 23(4), 471–493.
- Maulana, R., Helms-Lorenz, M., Ko, J., Chun, S., Lee, O., Irnidayanti, Y., De Jager, T., Coerzee, T. Koul, R. B., & Shahzad, A. H. (2019, April). Can we compare teaching behaviour across national contexts? Rasch modelling and differential item functioning approach. Poster presented at the 2019 American Educational Research Association Annual Meeting, Toronto, Canada.
- Ministry of Education of the People's Republic of China (MOE of China). (2011). *Curriculum Standard for Teacher Education (Trial)*. Retrieved from http://www.moe.gov.cn/srcsite/A10/s6991/201110/t20111008_145604.html
- Ministry of Education of the People's Republic of China (MOE of China). (2014). *Opinions on the implementation of outstanding teacher preparation plan*. Retrieved from http://www.moe.gov.cn/srcsite/A10/s7011/201408/t20140819_174307.html
- Moon, J. (2007). Getting the measure of reflection: Considering matters of definition and depth. *Journal of Radiotherapy in Practice*, 6(4), 191–200.
- O'Connell, C., Anderson, J., & Coe, E. (2000). Creative pathways to professional development. *Educational Developments*, 1, 8–10.
- Oner, D., & Adadan, E. (2011). Use of web-based portfolios as tools for reflection in preservice teacher education. *Journal of Teacher Education*, 62(5), 477–492.
- Pedro, J. Y. (2005). Reflection in teacher education: Exploring preservice teachers' meanings of reflective practice. *Reflective Practice*, 6(1), 49–66.

- Prilla, M., & Renner, B. (2014, November). Supporting collaborative reflection at work: A comparative case analysis. In *Proceedings of the 18th international conference on supporting group work* (pp. 182–193).
- Ryan, M., & Ryan, M. (2013). Theorising a model for teaching and assessing reflective learning in higher education. *Higher Education Research & Development*, 32(2), 244–257.
- Sammons, P., Kington, A., Lindorff-Vijayendran, A., & Ortega, L. (2014). *Inspiring teachers: Perspectives and practices*. CfBT Education Trust.
- Sammons, P., Lindorff, A. M., Ortega, L., & Kington, A. (2016). Inspiring teaching: Learning from exemplary practitioners. *Journal of Professional Capital and Community*, 1(2), 124–144.
- Schön, D. (1983). *The reflective practitioner. How professionals think in action*. Basic Books.
- Snoek, M. (2003). The use and methodology of scenario making. *European Journal of Teacher Education*, 26(1), 9–19.
- Tyrell, J., Lo, M.-L., Sankey, D., & Sam, C. A. (2013). *Field experience handbook (preservice education programmes)*. The Hong Kong Institute of Education.
- Van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research*, 49(2), 127–152.
- Van de Grift, W. J. (2014). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement*, 25(3), 295–311.
- Van der Lans, R. M., van de Grift, W. J., & van Veen, K. (2017). Individual differences in teacher development: An exploration of the applicability of a stage model to assess individual teachers. *Learning and Individual Differences*, 58, 46–55.
- Wang, Q., & Quek, C. L. (2015). *Investigating collaborative reflection with peers in an online learning environment*. International Society of the Learning Sciences, [ISLS].
- Wang, Y., Lei, J., & Fong, Y. (2020, April). *Analysing and evaluating preservice and novice teachers' reflective dialogues on teaching in a virtual learning environment* [Symposium]. AERA Annual Meeting San Francisco, CA.
- Williams, R., & Grudnoff, L. (2011). Making sense of reflection: A comparison of beginning and experienced teachers' perceptions of reflection for practice. *Reflective Practice*, 12(3), 281–291.

Ye Wang is a dual doctoral candidate of The Education University of Hong Kong and Hiroshima University. Her research interests are teacher development and teaching quality of preservice teachers, especially in the fields of teaching reflection, classroom observation, and teacher efficacy. She has contributed to a number of international research projects in cooperation with the Netherlands, the United States, Japan, and South Korea.

Dr. James Ko is an Associate Professor at the Department of Policy Leadership and Co-Director of the Joseph Lau Luen Hung Charitable Trust Asia Pacific Centre for Leadership and Change at the Education University of Hong Kong. Before his doctoral study, James was an EFL teacher for about 20 years and led two functional teams in a secondary school for 10 years. He is a recurrent grantee of the RGC and UGC grants and the principal investigator of 23 projects, collaborating with local academics and overseas researchers on 40 projects. He has supervised 14 doctoral students with 8 completed.

Peng Wang is the chief physician in the Orthopedics Department, Affiliated Hospital of Hebei University of Chinese Medicine, China. He is a regular member of the Committee of Minimal Invasive Spine Surgery (MISS) in China. As an expert in teacher training, he is often invited to give demo courses for teachers and students in Chinese medicine universities and colleges all over China. He serves in the committee of the 17th National Teaching Competition of Chinese Medicine Universities in 2019.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 26

Effective and Inspiring Teaching in STEM Classrooms: Evidence from Classroom Observations with Instrument Comparisons



James Ko 

Abstract This book chapter reports findings in a case study on the video clips of 97 STEM lessons at a local secondary school. The impact of Effective and inspiring teaching on student engagement in classrooms was explored using the same high-inference classroom observation instruments. Cluster analysis indicated that effective teaching dimensions tended to cluster together. However, inspiring teaching dimensions (i.e., Flexibility, Innovative teaching, and Teaching reflective thinking) tended to cluster with Teaching collaborative learning. While there was no subject difference for inspiring teaching practices, Mathematics significantly performed the best and Technology the worst in effective teaching practices. Multiple regression results indicated that both effective and inspiring teaching practices have a significant but moderate impact on learner engagement, but none showed significant effects on student engagement. In contrast, while the effective teaching dimension Professional knowledge and expectations positively affected overall teaching quality perceptions.

Keywords Effective teaching · Inspiring teaching · Instrument comparison · Student engagement

1 Introduction

This study represents a classroom observation approach to capture rich information about classroom behaviours and activities through instruments developed to observe generic teaching behaviours across subjects, grades, and contexts. The research comparing effective and inspiring teaching is justified because the conceptual boundary between effective teaching and inspiring teaching was unclear, indicating

J. Ko (✉)

Department of Education Policy and Leadership, Education University of Hong Kong,
Tai Po, Hong Kong
e-mail: jamesko@eduhk.hk

a theoretical overlap (Sammons et al., 2014). An instrument comparison was adopted as a methodology strategy because of previous international studies (e.g., Kington et al., 2014; Kane & Stagier, 2012; Kane et al., 2011; Sammons et al., 2014). Another methodology strategy was to limit the lesson sample to a single school. This strategy was also adopted in the previous projects to illuminate the rich variations across departments within a school suggested by Sammons et al. (1997).

2 Theoretical Background

This study extended empirical works on teacher effectiveness (Kington et al., 2014; Ko et al., 2015; Ko et al., 2016; Sammons et al., 2014). The following sections examine three interrelated issues: the comparisons between effective and inspiring teaching, classroom observations with high-inference instruments, and contextual influences on teaching quality variations.

2.1 *Characteristics of Inspiring Teachers and Relations with Effective Teaching*

Compared to the vast amount of literature on teacher effectiveness (see Ko & Sammons, 2013; Hattie, 2009), inspiring teaching is minimal. Harmin and Toth (2006, p.16) outlined some professional characteristics of inspiring teachers, but they suggested what these teachers might do in the classroom. Inspiring teachers may make a lesson more inspiring through four steps: *targeting* (i.e., “maintain clear standards for themselves with a strong sense of their ideals and directions”), *adjusting* (i.e., “able to adjust their teaching when they choose to do so and not reluctant to explore something new if they sense it might help them better serve their ideals”), *balancing* (i.e., “maintain a fair measure of personal balance in their work”), and *supporting* (i.e., “willing to share ideas and talk with colleagues about professional questions, including their personal confusions and weaknesses.”

In England, Sammons et al. (2014) conducted a study to explore inspiring teaching and found that inspiring teachers shared many effective teachers’ characteristics. Based on the lesson observations of 17 inspiring primary and secondary teachers, Sammons et al. (2016, p. 136) found many practices and behaviours typically associated with highly effective teaching included:

- Creating a positive, safe, and supportive climate for learning
- Managing behaviour, space, time, and resources efficiently and effectively
- Implementing clear instruction, including explicit and high expectations and objectives for learning
- Demonstrating good behaviour management skills and efficient use of learning time

- Skilful use of questioning and feedback to make lessons highly interactive and extend learning.

In addition, inspiring teachers were found to be:

- Using largely informal approaches to meet individual student needs
- Promoting high levels of student engagement and motivation through varied learning activities and arrangements
- Seeking ways to promote and honour student choice and input
- Using a wide variety of activities or approaches throughout a lesson
- Showing high levels of commitment and care for students' learning and well being
- Developing and reinforcing positive relationships with students

Sammons et al. (2014, p.16) pointed out that their participant teachers considered that “inspiring and being effective were two related and mutually-dependent aspects of teaching” such that “being inspiring was much more due to the link with relationships.” For example, like their effective colleagues, inspiring teachers can also develop a positive relationship with students, making their lessons more enjoyable, stimulating and engaging (Sammons et al., 2014). Effective teachers can make their lessons engaging through better structuring and stronger connections between the learning activities with students' daily experiences (Ko et al., 2015). However, inspiring teachers seem to achieve similar influences on students through stronger personal connections with students, simultaneously mixing well three aspects of teaching *Positive classroom management, Enthusiasm for teaching, and Positive relationships with children* (Fig. 26.1).



Fig. 26.1 Characteristics of inspiring teachers in Sammons et al. (2014)

Sammons et al. (2014, 2016) did not develop any instrument to distinguish the classroom practices of inspiring teachers. Instead, they used the same instruments used in Kington et al., (2014), which are more appropriate to capture effective teachers' generic teaching characteristics. Motivated to address the lack of a valid classroom observation instrument to measure and characterise the similarities and differences between effective and inspiring teaching quantitatively, Ko et al. (2016, 2019a, b) conceptualised three aspects of teaching behaviours in Fig. 26.1 more explicitly related to inspiring teaching: *Innovative Teaching, Flexibility, Reflectiveness and Collaboration*. For example, inspiring teachers are often more willing to develop stronger collaborations and offer more support than colleagues than other teachers (Sammons et al., 2014). International research by OECD indicated that teacher collaboration helps support teacher reflection and thus forms an essential feature of professional practice (Vieluf et al., 2012). Pedagogical innovations are also strongly associated with teachers' reflections through professional collaborations with other teachers (Vieluf et al., 2012).

Based on a secondary analysis of 206 lesson videos selected from 306 Hong Kong lessons of the 538 lessons by Ko et al. (2015), Ko et al. (2016, 2019a, b) identified two clusters in hierarchical cluster analysis results. Cluster 1 with eight factors represents *Effective Teaching: Enthusiasm for teaching, Positive relationships with students, Purposeful and relevant teaching, Safe classroom climate, Stimulating learning environment, Positive classroom management, Assessment for learning, and Professional knowledge and expectations*. Cluster 2 indicates *Inspiring Teaching* with factors: *Flexibility, Teaching reflective thinking, and Innovative teaching*.

2.2 Classroom Observation Using High-Inference Instruments

High-inference classroom observation instruments are often preferable in classroom research. While high-inference instruments are generally more subjective, they are much more cost-effective to conduct than low-inference instruments. High-inference instruments require the observer to make high inferences or judgements about the behaviours and their impacts observed in the classroom (Muijs & Reynolds, 2017; O'Leary, 2020; Schaffer et al., 1994).

Among the two low-inference and three high-inference instruments that Ko et al. (2015) compared, the *International Comparative Analysis of Teaching and Learning (ICALT)* (formerly known as the *Quality of Teaching Scale*; van de Grift 2007, van de Grift et al., 2014) were found distinguishing effective teaching behaviours more clearly. By conducting secondary data analysis of the same set of video-recorded lessons with a similar high-inference observation instrument specifically for measuring inspiring teaching, Ko et al. (2016, 2019a, b) developed a new high-inference instrument to compare effective and inspiring teaching with the generic behavioural characteristics of effective teachers characterised in the *ICALT*. Thus, this study can extend Ko et al.'s (2016, 2019a, b) work to examine effective and

inspiring teaching in STEM subjects of a school, which is presumably a more confined context.

2.3 Contextual Influences on Variations of Teaching Quality

In the literature, while variations across schools in an education system are often the focus of school effectiveness research, Sammons et al. (1997) showed that within-school variations were often more extensive than between-school variations. Effective departments exist in ineffective schools, while effective departments also exist in effective schools. Ko (2010) noted that considerable variations existed in the same teachers' different classrooms because teaching consistency is hard to maintain teaching effectiveness or some teachers who seemed to struggle with teaching specific student groups like students with special needs or affected by the washback effect of the public examination.

An empirical work by Opdenakker and Van Damme (2007) suggested significant influences of school context, student composition and school leadership on school practice and outcomes in secondary education. Contextual effects on effective teaching were inconclusive (Ko et al., 2015). While no city showed a dominance of effective or less effective teachers, considerable differences in the school sector, subject, and location contrasts were evident (Ko et al., 2015). Interestingly, the teaching effectiveness patterns of highly effective and highly ineffective teachers in different cities look alike. Studies in China (e.g., Li, 2015; Walker et al., 2012) showed the increasingly significant role of school principals in China in promoting schools' pedagogical innovations. Chinese teachers also participated in professional development and led research to enhance teaching and learning more often than Hong Kong teachers. These results suggest that we need to develop an appropriate interview protocol that goes beyond investigating the teaching practices of Hong Kong and Guangzhou schools but looks into the impact of broader educational contexts and the characteristics within schools such as leadership, instructional management, department and school policies.

2.4 Research Questions

To explore the overlapping relationships between effective and inspiring teaching, I continued to adopt the instrument strategy in addressing the following research questions:

1. What specific teaching behaviours/dimensions can be characterised as inspiring in the observed STEM classrooms?
2. Are there differences in teaching quality among different STEM subjects?

3. Do effective teaching and inspiring teaching impact student engagement?
4. Do effective teaching and inspiring teaching impact perceptions of overall teaching quality?

3 Methods

3.1 Samples

As a case study of a single local English medium secondary school in Hong Kong, the lesson video sample consisted of 97 lessons in four STEM subjects: Mathematics, Science, Technology, and Personal, Social and Health Education (PSHE). The academic attainment of the school is about the top one-third. Despite a tuition fee of about HK3000 per month, the subscription is keen among families with middle socio-economic backgrounds in the district. The school initially videotaped all lessons for internal teacher evaluation and professional development purposes. Ethical consent forms were obtained through the school administration.

3.2 Instruments

The two classroom observation instruments employed in this study were the same as those in Ko et al. (2016, 2019a, b). Both instruments were high-inference by nature, requiring the subjective judgements of the raters. ICALT was well established and validated across many countries (Maulana et al., 2020), but CETIT also has high reliability and validity (Ko, Sammons & Kyriakides, 2016).

3.2.1 International Comparative Analysis of Teaching and Learning (ICALT)

Originated as an instrument for inspections, the *International Comparative Analysis of Learning and Teaching (ICALT)* (van de Grift, 2014) was an instrument that combined low and high-inference components. Raters have to indicate the absence or presence of teaching behaviours associated before rating the teacher performance along 32 teaching indicators to determine their strengths on a 4-point scale, from 'mostly weak' to 'mostly strong'. As depicted in Table 26.1, these teaching indicators are theoretically grouped further into six domains: *Safe and stimulating learning climate*, *Efficient organisation*, *Clear and structured instructions*, *Intensive and activating teaching*, *Adjusting instructions for learner differences*, and *Teaching learning strategies*. For the ease of associating teaching behaviours with student engagement during classroom observations, the ICALT also contained a three-item (e.g., "...take an active approach to learn") domain to document learner engagement.

Table 26.1 Teacher dimensions, no of items in each dimension, and item examples of ICALT

Teaching dimension	Items	Item examples
<i>Safe and stimulating learning climate</i>	4	The teacher shows respect for learners in his/her behaviour and language
		The teacher maintains a relaxed and friendly classroom atmosphere
<i>Efficient organisation</i>	4	The teacher ensures the lesson proceeded in an orderly manner
		The teacher monitors to ensure learners carry out activities in the appropriate manner
<i>Clear and structured instructions</i>	7	The teacher presents and explains the subject material in a clear manner
		The teacher gives feedback to learners
<i>Intensive and activating teaching</i>	7	The teacher offers activities and work forms that stimulate learners to take an active approach
		The teacher stimulates the building of self-confidence in weaker learners
<i>Adjusting instructions for learner diversity</i>	4	The teacher evaluates whether the lesson aims have been reached
		The teacher offers weaker learners extra study and instruction time
<i>Teaching learning strategies</i>	6	The teacher teaches learners how to simplify complex problems
		The teacher teaches learners to check solutions
<i>Indicator for the learners</i>		
<i>Learner engagement</i>	3	... are fully engaged in the lesson
		... show that they are interested
		... take an active approach to learning

3.2.2 Comparative Analysis of Effective Teaching and Inspiring Teaching (CETIT)

Ko et al. (2016) used the Delphi method to finalise 68 items and validated a new high-inference classroom observation instrument with 12 teaching aspects of effective and inspiring teaching behaviours. Ten of the 12 aspects were identified qualitatively by Sammons et al. (2014). Ko et al. (2016) hypothesised that *Flexibility, Teaching reflective thinking, Innovative teaching, and Teaching collaborative learning*. Respective examples of teaching behaviours were “*The teacher allowed options for students in their seatwork,*” “*The teacher asked students to comment on his/her viewpoint,*” “*The teacher used ICT in teaching,*” “*The teacher told students how to share their work in a task.*”

Reflectiveness and collaboration were considered characteristics of inspiring teachers in Sammons et al.’s (2014) study. However, Ko et al. (2016) considered inspiring teachers to promote collaborative learning and develop students’ reflective thinking as two distinctive classroom practices. They also distinguished a safe and stimulating classroom climate as they could be conceptually and empirically

Table 26.2 Teacher dimensions, no of items in each dimension, and item examples of CETIT

Teaching dimension	Items	Item examples
<i>Enthusiasm for teaching</i>	5	The teacher made learning exciting for most students.
<i>Positive relationships with students</i>	6	The teacher knew the students well individually.
<i>Purposeful and relevant teaching</i>	6	The teacher encouraged students to find different solutions for a problem.
<i>Safe classroom climate</i>	5	The teacher promoted learners' self-confidence.
<i>Stimulating learning environment</i>	4	The teacher asked students to think about how to link up what they learned.
<i>Positive classroom management</i>	4	The teacher used the time for learning efficiently.
<i>Assessment for learning</i>	8	The teacher provided appropriate feedback to students.
<i>Professional knowledge and expectations'</i>	5	The teacher showed clear concepts in his/her classroom language.
<i>Flexibility</i>	5	The teacher allowed options for students in their own seatwork.
<i>Teaching reflective thinking</i>	9	The teacher asked students to comment on his/her viewpoint.
<i>Innovative teaching</i>	5	The teacher used ICT in teaching.
<i>Teaching collaborative learning</i>	5	The teacher told students how to share their work in a task.

different in some studies (e.g., van de Grift, 2007). *Assessment for learning* and *Professional knowledge and expectations* were not studied previously (Kyriakides & Creemers, 2008; Day et al., 2008; Ko et al., 2015) but were included for their potential to extend the existing models of teacher effectiveness empirically (Table 26.2).

3.3 Raters

Four research assistants with varied research experience in classroom observation observed the lesson videos after calibrations with training videos and two lesson videos in the sample. They had to discuss the discrepancies in evaluations. Experience, training and calibration were crucial for achieving high reliability. Inter-rater reliability of .79 was achieved before they started to do observation independently.

4 Results

4.1 Descriptive Statistics

Table 26.3 summarises the mean, standard deviation, and reliability of each teaching dimension of the two instruments, CETIT and ICALT. It is not surprising that *Positive classroom management*, *Safe classroom climate*, and *Safe and stimulating*

Table 26.3 Mean, variance, and reliability (Cronbach alpha) of each teaching dimension

Domain	Dimension	Mean	S. D.	Cronbach Alpha
CETIT	Based on a 5-point scale			
Effective	Positive relationships with students	2.99	.76	.78
Effective	Purposeful and relevant teaching	3.23	.88	.81
Effective	Safe classroom climate	3.95	.67	.76
Effective	Stimulating learning environment	2.89	.81	.76
Effective	Positive classroom management	4.21	.71	.90
Effective	Assessment for learning	2.96	.62	.71
Effective	Professional knowledge and expectations	3.80	.58	.75
Effective	Enthusiasm for teaching	3.69	.80	.89
Inspiring	Flexibility	1.56	.63	.92
Inspiring	Innovative teaching	3.24	.97	.71
Inspiring	Teaching reflective thinking	2.13	.59	.71
Mixed	Teaching collaborative learning	2.75	1.14	.83
ICALT	Based on a 4-point scale			
Effective	Safe and stimulating learning climate	3.39	.57	.86
Effective	Effective organisation	3.18	.57	.90
Effective	Clear, structured instruction	2.83	.60	.93
Effective	Intensive and activated teaching	2.48	.54	.84
Effective	Adjusted instruction for catering learner diversity	1.61	.38	.41
Effective	Teaching learning strategy	1.87	.67	.86
Effective	Learner engagement	2.89	.57	.86
Effective	The judgement of overall teaching quality	2.94	.41	

learning climate have the highest means, while *Flexibility*, *Adjusted instruction for catering to learner diversity*, and *Teaching learning strategy* have the lowest means. Because in line with the research literature, these dimensions represent the most straightforward and most challenging aspects of teaching. Most standard deviations are not high, except for *Teaching collaborative learning*. Most teaching dimensions' reliability scores were well above .7, ranging from .7 to .93, indicating good reliability except for *Adjusted instruction for catering to learner diversity*, which has an alpha of .41, below the acceptable reliability of .7 for a scale in education research (Taber, 2017).

Table 26.4 summarises the two-tailed Pearson correlations between the CETIT dimensions and the ICALT dimension *Learner engagement* and the judgement of *Overall teaching quality*. Among all teaching dimensions, *Flexibility* and *Innovative teaching* are least likely to be associated with other teaching dimensions, including *Professional knowledge and expectations*, suggesting inspiring teaching practices do not necessarily require professional solid content knowledge. However, *Flexibility* is correlated significantly with *Teaching reflective thinking* and *Positive relationships with students*, suggesting teaching students reflective thinking may require some flexibility (or 'thinking out of the box' attitude) and reflects positive relationships with students. Teachers sometimes may have to be flexible for

Table 26.4 Pearson correlations of teaching dimensions of CETIT and ICALT

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
<i>I</i> Positive relationships with students		.543**	.471**	.484**	.252*	.479**	.623**	.491**	.341**	-.121	.548**	.335**	.560**	.237*	.395**	.537**	.108	.265**	.185	.442**
<i>2</i> Purposeful and relevant teaching	.543**		.646**	.805**	.448**	.694**	.598**	.473**	.216*	.253*	.477**	.228*	.541**	.556**	.742**	.777**	.436**	.700**	.431**	.632**
<i>3</i> Safe classroom climate	.471**	.646**		.633**	.681**	.612**	.707**	.479**	.075	.254*	.384**	.366**	.684**	.665**	.590**	.646**	.437**	.460**	.452**	.672**
<i>4</i> Stimulating learning environment	.484**	.805**	.633**		.409**	.667**	.446**	.298**	.173	.252*	.494**	.270**	.499**	.537**	.600**	.699**	.495**	.649**	.441**	.558**
<i>5</i> Positive classroom management	.252*	.448**	.681**	.409**		.471**	.723**	.544**	-.238**	.120	.207*	.080	.461**	.803**	.649**	.499**	.301**	.422**	.466**	.680**
<i>6</i> Assessment for learning	.479**	.694**	.612**	.667**	.471**		.590**	.376**	.203*	.087	.532**	.273**	.543**	.473**	.635**	.701**	.460**	.528**	.359**	.586**
<i>7</i> Professional knowledge and expectations	.623**	.598**	.707**	.446**	.723**	.590**		.756**	.131	.034	.473**	.170	.595**	.635**	.630**	.611**	.313**	.377**	.407**	.731**
<i>8</i> Enthusiasm for teaching	.491**	.473**	.479**	.298**	.544**	.376**	.756**		.185	.066	.316**	.079	.469**	.502**	.524**	.437**	.159	.158	.219*	.504**
<i>9</i> Flexibility	.341**	.216*	.075	.173	-.238**	.203*	.131	.185		.107	.479**	.217*	.064	-.201*	-.088	.113	.086	-.145	-.182	-.111
<i>10</i> Innovative teaching	-.121	.253*	.254*	.252*	.120	.087	.034	.107	.066		.086	.128	.117	.216*	.226*	.156	.253*	.170	.128	.083
<i>11</i> Teaching reflective thinking	.548**	.477**	.384**	.494**	.207*	.532**	.473**	.316**	.479**	.086		.366**	.395**	.170	.293**	.522**	.230*	.274**	.157	.383**

12	Teaching collaborative learning	.335**	.228*	.366**	.270**	.080	.273**	.170	.079	.217*	.128	.366**	1	.428**	.058	.015	.375**	.240*	.051	.147	.294**
13	Safe and stimulating learning climate	.560**	.541**	.684**	.499**	.461**	.543**	.595**	.469**	.064	.117	.395**	.428**	1	.466**	.518**	.695**	.148	.255*	.359**	.526**
14	Effective organisation	.237*	.556**	.665**	.537**	.803**	.473**	.635**	.502**	-.201*	.216*	.170	.058	.466**	1	.791**	.583**	.439**	.552**	.606**	.752**
15	Clear structured instruction	.395**	.742**	.590**	.600**	.649**	.635**	.630**	.524**	-.088	.226*	.293**	.015	.518**	.791**	1	.737**	.462**	.651**	.503**	.677**
16	Intensive and activated teaching	.537**	.777**	.646**	.699**	.499**	.701**	.611**	.437**	.113	.156	.522**	.375**	.695**	.583**	.737**	1	.461**	.640**	.515**	.707**
17	Adjusted instruction for catering learner diversity	.108	.436**	.437**	.495**	.301**	.460**	.313**	.159	.086	.253*	.230*	.240*	.148	.439**	.462**	.461**	1	.536**	.400**	.407**
18	Teaching learning strategy	.265**	.700**	.460**	.649**	.422**	.528**	.377**	.158	-.145	.170	.274**	.051	.255*	.552**	.651**	.640**	.536**	1	.474**	.552**
19	Learner engagement	.185	.431**	.452**	.441**	.466**	.359**	.407**	.210*	-.182	.128	.157	.147	.359**	.600**	.503**	.515**	.400**	.474**	1	.733**
20	The judgement of overall teaching quality	.442**	.632**	.672**	.558**	.680**	.586**	.731**	.504**	-.111	.083	.383**	.294**	.526**	.752**	.677**	.707**	.407**	.552**	.733**	1

*, Correlation is significant at the .01 level (2-tailed)

**, Correlation is significant at the .05 level (2-tailed)

teaching students collaborative learning or assessing them for learning. The cost of being flexible would be an impression of ‘poor’ classroom management or ‘ineffective’ organisation, as indicated by the significant negative correlations between these teaching dimensions.

4.2 *Categorisation of Effective and Inspiring Teaching Behaviours*

We decided the clusters using average linkage to estimate the distance among factors, which overcomes the shortcoming of single and complete linkage (Yim & Ramdeen, 2015). Based on the Agglomeration Coefficients of hierarchical cluster analysis conducted with SPSS version 24, the results suggested the clustering process should stop or stay at stage 8 (Table 26.5. and Fig. 26.2). By stopping the clustering at this point, the factors were clustering into four categories (Fig. 26.3).

Hence, grouping dimensions Safe classroom climate, Professional knowledge and expectations, Positive classroom management, Enthusiasm for teaching, Purposeful and relevant teaching, Stimulating learning environment, Assessment for learning and Positive relationship with students formed the first cluster. Dimensions *Flexibility* and *Reflectiveness* were grouped as the second cluster. Innovative teaching and Teaching collaborative learning were two isolated clusters.

4.3 *Differential Teaching Behaviours Among STEM Subjects*

As depicted in Table 26.6, the means between the two instruments seemed to show similar patterns since the Mathematics lessons had the highest means. In contrast, regardless of instruments, Technology lessons had the lowest, except for that the

Table 26.5. Agglomeration schedule of hierarchical cluster analysis

Stage	Cluster combined		Coefficients	Stage cluster first appears		Next stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	4	11	.930	0	0	2
2	4	6	1.330	1	0	5
3	3	5	1.525	0	0	4
4	3	10	2.027	3	0	6
5	4	12	2.286	2	0	8
6	1	3	3.022	0	4	8
7	2	7	3.542	0	0	10
8	1	4	4.215	6	5	9
9	1	9	6.347	8	0	11
10	2	8	8.190	7	0	11
11	1	2	8.616	9	10	0

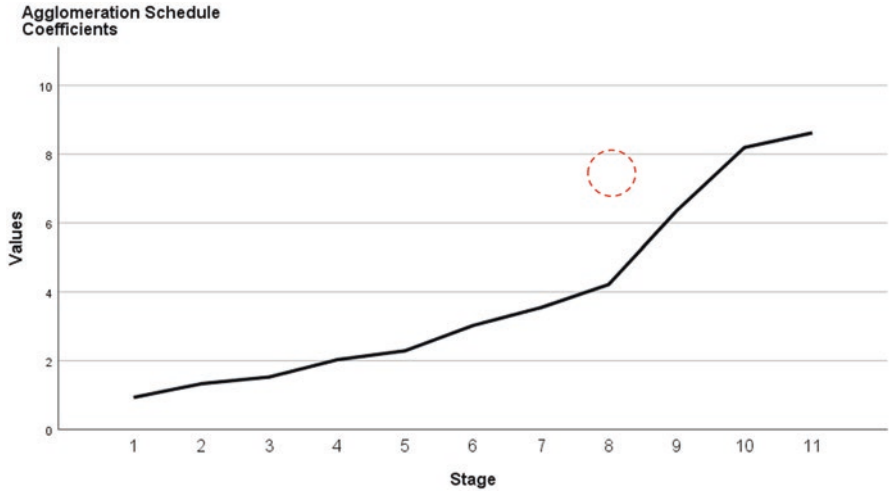


Fig. 26.2 Agglomeration Schedule of hierarchical cluster analysis

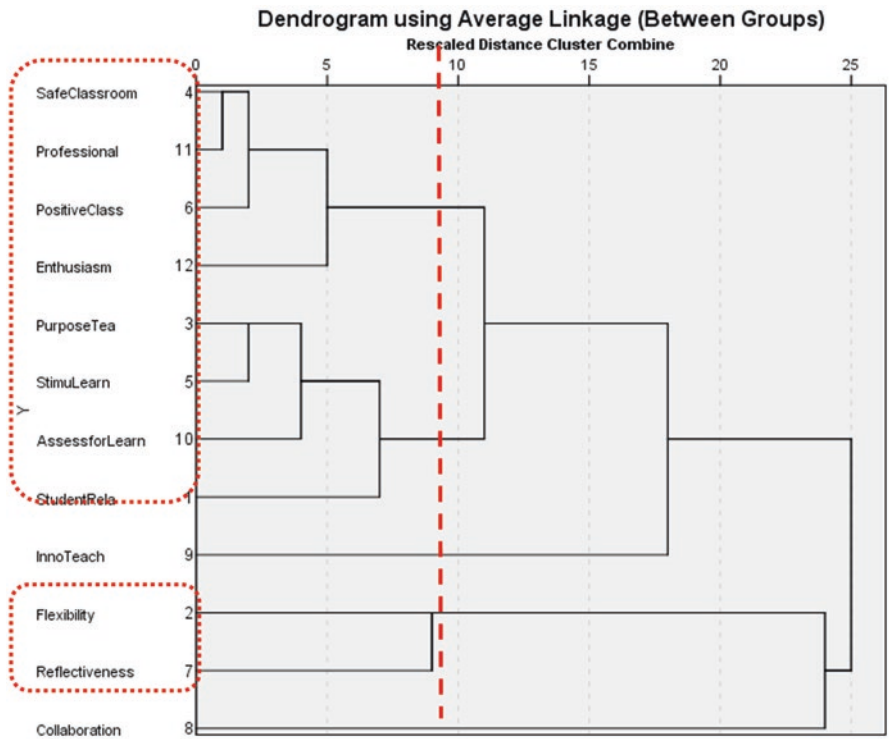


Fig. 26.3 Clustering of effective and inspiring behaviours

Table 26.6 Comparisons of STEM subjects with means and standard deviations

		N	Mean	Std. Deviation	Std. Error	Minimum	Maximum	Between-Component Variance	
ICALT mean	Mathematics	27	2.875	.331	.064	2.31	3.45		
	PSHE	21	2.510	.289	.063	1.88	3.13		
	Science	26	2.558	.364	.071	1.46	3.27		
	Technology	23	2.234	.492	.103	1.07	2.90		
	Total	97	2.559	.437	.044	1.07	3.45		
	Model	Fixed effects			.376	.038			
		Random effects				.134			.065
CETIT Effective teaching mean	Mathematics	27	3.713	.431	.083	2.83	4.43		
	PSHE	21	3.522	.375	.082	2.71	4.18		
	Science	26	3.457	.485	.095	1.61	4.21		
	Technology	23	3.129	.765	.160	1.00	3.91		
	Total	97	3.464	.566	.057	1.00	4.43		
	Model	Fixed effects			.533	.054			
		Random effects				.123			.048
CETIT Inspiring teaching Mean	Mathematics	27	2.351	.414	.080	1.67	3.01		
	PSHE	21	2.201	.496	.108	1.30	3.17		
	Science	26	2.220	.545	.107	1.47	3.46		
	Technology	23	2.461	.537	.112	1.22	3.40		
	Total	97	2.309	.502	.051	1.22	3.46		
	Model	Fixed effects			.499	.051			
		Random effects				.059			.004
CETIT mean	Mathematics	27	3.294	.379	.073	2.47	3.92		
	PSHE	21	3.113	.426	.093	2.21	3.78		
	Science	26	3.100	.422	.083	1.63	3.79		
	Technology	23	2.928	.634	.132	1.18	3.57		
	Total	97	3.116	.482	.049	1.18	3.92		
	Model	Fixed effects			.471	.048			
		Random effects				.076			.014

ICALT average for Science lessons was higher than that for PSHE lessons, but vice versa for the average for the CETIT-effective teaching component. There was no subject difference for the CETIT ($F(3,96) = 2.522, p = .063$). However, there is a significant difference in the ICALT among four subjects ($F(3,96) = 12.18, p < .001$).

However, when instrument comparisons were narrowed down into details where the effective teaching component and the inspiring component of the CETIT were separate, the results showed exciting distinctions. First, there was a significant difference in the CETIT-effective teaching component among four subjects ($F(3,96) = 5.08, p < .001$). Second, while the CETIT-inspiring teaching component's variations remained insignificant ($F(3,96) = 1.37, p = .256$), Technology lessons had the highest mean because more innovative teaching was found in this subject. These results suggested that while both the ICALT and the CETIT-effective teaching component could distinguish the teaching quality of four subjects, the CETIT showed more variations at the teaching dimension level.

4.4 Impact of Effective Teaching and Inspiring Teaching on Student Engagement

Multiple regression analysis in SPSS version 24 was performed to explore the relative significance of the effective and inspiring teaching dimensions of CETIT in predicting student engagement. Learner engagement of ICALT was used as the dependent variable. The eight theoretical dimensions of effective teaching were entered first, followed by the four inspiring teaching dimensions to test the hierarchical models. Effective teaching practices had a significant but moderate impact ($R^2 = .32$ for Model 1, $p < .001$, effective teaching practices alone), but additional inspiring teaching component had an insignificant impact on learner engagement ($F = 1.174, p = .328$ for Model 2, both effective and inspiring teaching practices) (Table 26.7). None of the individual teaching dimensions was found to impact student engagement significantly. As results indicated that the basic constant model was significant, other factors such as subject differences might affect student engagement. As there were many variables in building both models, multicollinearity might have also affected the modelling results.

4.5 Impact Effective Teaching and Inspiring Teaching on the Overall Perception of Teaching Quality

Contrary to the results showing no significant impact of individual teaching dimensions on student engagement, models in Table 26.8 indicated significant effects of effective and inspiring teaching dimensions. While *Positive classroom management* ($\beta = .258$) and *Professional knowledge and expectations* strongly affected student engagement positively, the latter's strength was stronger ($\beta = .42$) ($R^2 = .652, F(8, 96) = 1.594, p < .001$ for Model 1, effective teaching component only). However, when inspiring teaching dimensions were added as predictors ($R^2 = .722, F(12, 96) = 17.305, p < .001$ for Model 2, both effective and inspiring teaching components),

Table 26.7 Regression model summary of teaching dimensions of CETIT as predictors^a

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change statistics					Durbin-Watson
					R Square change	F change	df1	df2	Sig. F change	
1	.569 ^b	.324	.262	.490	.324	5.261	8	88	.000	
2	.599 ^c	.359	.268	.488	.036	1.174	4	84	.328	2.134

a. Dependent Variable: Learner Engagement

b. Predictors: (Constant), *Effective teaching dimensions*: Professional knowledge and expectations, Stimulating learning environment, Positive relationships with students, Assessment for learning, Positive classroom management Safe classroom climate, Purposeful and relevant teaching, Enthusiasm for teaching

c. Predictors: (Constant), *Effective teaching dimensions*: Professional knowledge and expectations, Stimulating learning environment, Positive relationships with students, Assessment for learning, Positive classroom management Safe classroom climate, Purposeful and relevant teaching, Enthusiasm for teaching; *Inspiring teaching dimensions*: Innovative Teaching, Flexibility, Teaching reflective thinking, Teaching collaborative learning

Professional knowledge and expectations ($\beta = -.482$) remained significantly affecting overall teaching quality perceptions. Inspiring teaching dimensions *Flexibility* and *Teaching collaborative learning* ($\beta = .178$) affected perceptions of overall teaching quality. Interestingly, more teaching flexibility was perceived negatively ($\beta = -.251$). Again, results indicated that the basic constant model was significant, suggesting other factors (such as subject differences may affect judgments of teaching quality).

5 Discussions

Overall the study results showed that effective teaching in the two instruments looked similar but differed much from inspiring teaching. The former indicates more innovative and require flexibility in application, while the latter may be more generic. Both correlation and clustering results indicated that teaching flexibility is associated with teaching students reflective thinking, and they may also be indispensable for innovative teaching and collaborative learning. Inspiring teachers may encourage students to reflect on their own and others' views and engage them in collaborative learning activities. Thus, it seems that flexibility is a teaching asset not necessarily co-occurring as effective teaching practices.

Interestingly, correlations indicated that strong professional knowledge might hinder the adoption of innovative teaching and hamper teaching flexibility. Inspiring teaching may emerge in the early teaching stage when a teacher still has not shown exceptionally strong in his/her professional knowledge. Perhaps some professional development programs can support teachers with sound professional knowledge to adopt more innovative and flexible teaching. The following sessions address the limitations, significances, implications for professional development and conclusion.

Table 26.8 Regression model summary of effects of CETIT teaching dimensions^a on overall teaching quality

Model		Unstandardised Coefficients		Standardised Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.785	.184		4.255	.000
	Positive relationships with students	-.114	.098	-.152	-1.162	.248
	Purposeful and relevant teaching	.071	.111	.110	.643	.522
	Safe classroom climate	.052	.130	.061	.398	.692
	Stimulating learning environment	.185	.117	.265	1.581	.117
	Positive classroom management	.183	.121	.229	1.515	.133
	Assessment for learning	-.044	.124	-.048	-.353	.725
	Professional knowledge and expectations	.136	.185	.138	.735	.464
2	(Constant)	1.348	.408		3.303	.001
	Positive relationships with students	-.078	.105	-.104	-.743	.459
	Purposeful and relevant teaching	.096	.113	.147	.842	.402
	Safe classroom climate	.047	.134	.055	.350	.727
	Stimulating learning environment	.183	.119	.261	1.534	.129
	Positive classroom management	.077	.135	.096	.571	.569
	Assessment for learning	-.030	.129	-.032	-.230	.818
	Professional knowledge and expectations	.318	.213	.322	1.488	.140
	Flexibility	-.187	.106	-.207	-1.772	.080
	Teaching reflective thinking	-.017	.122	-.018	-.142	.887
	Innovative teaching	.007	.060	.012	.122	.903
	Enthusiasm for teaching	-.103	.101	-.144	-1.023	.309

a. Dependent Variable: Final Judgement of Overall Teaching Quality

b. Predictors: (Constant), *Effective teaching dimensions*: Professional knowledge and expectations, Stimulating learning environment, Positive relationships with students, Assessment for learning, Positive classroom management Safe classroom climate, Purposeful and relevant teaching, Enthusiasm for teaching

c. Predictors: (Constant), *Effective teaching dimensions*: Professional knowledge and expectations, Stimulating learning environment, Positive relationships with students, Assessment for learning, Positive classroom management Safe classroom climate, Purposeful and relevant teaching, Enthusiasm for teaching; *Inspiring teaching dimensions*: Innovative Teaching, Flexibility, Teaching reflective thinking, Teaching collaborative learning

5.1 *Distinctions Between Effective and Inspiring Teaching*

Empirical studies on the distinctions between effective and inspiring teaching are rare because we lack proper theoretical frameworks and associated instruments to distinguish them. Sammons and her colleagues (2014, 2016) contended that an important distinction between inspiring and effective teaching lies in our theory and methodology as well as our perspective of measurement and evaluation. Sammons et al. (2016) argued that theories without direct observation and measurement, but primarily on attitudinal measures, interviews, and similar indirect measures, are inadequate. Regarding teacher evaluation, as “the word ‘inspiring’ casts a wider net linking with affective and social-behavioural outcomes, [this] raises questions about the extent to which inspirational outcomes overlap with effective outcomes, and whether effectiveness is compatible with, part of, or different from inspiring practice” (Sammons et al., 2016, p. 125).

Similar to findings on English and Mathematics in Ko et al. (2019a, b), the cluster analysis supported a distinction of effective and inspiring teaching. However, only two of the teaching dimensions originally proposed as inspiring teaching in Ko et al. (2015), that is, *Flexibility* and *Reflectiveness* or *Teaching reflective thinking*. This raises the question that some aspects are basic or occur in a broader range of classrooms, while some are more context or subject-specific. Moreover, while Sammons et al. (2014, 2016) suggested that inspiring teachers were “dedicated, positive, and caring” teachers in their study, conceptually related factors like *Enthusiasm for teaching*,

Positive relationships with students, *Safe classroom climate*, and *Positive classroom management* were associated with other factors associated with effective teaching instead. We are not sure whether the different results might involve cultural influences. That is, effective teachers in Hong Kong samples were more dedicated, positive, and caring. Though it is hard to conceive that inspiring teachers do not have these characteristics, our current study cannot provide conclusive answers.

5.2 *Innovative Teaching in Inspiring Teaching and Professional Development Implications*

Our clustering results indicated that innovative teaching did not associate closer with inspiring teaching as one might expect. In the current conceptualisation, the factor *Innovative Teaching* concerns the extent to which ICT is applied in teaching and learning, which could be a narrow conception of innovativeness for other researchers. For example, Maass et al. (2019) consider that innovative teaching approaches also include those that can combine and scale-up material- and community-based implementation strategies. In OECD’s (2014) articulation, innovative teaching can concern regrouping educators and teachers for collaborative planning, orchestration and professional development, team teaching to target

specific groups of learners, widening pedagogical repertoires like inquiry learning, authentic learning, and mixes of pedagogies, while pedagogical possibilities in ‘technology-rich’ environment are just a few narrower options. This may imply that the current conceptualisation of innovative teaching is too restrictive to include teaching practices that can be connected to inspiring teaching.

Moreover, innovative teaching is still a weaker aspect in non-technology STEM subjects, perhaps in other academic subjects too. This is a little surprising that subjects that are traditionally conceptualised as STEM subjects like Mathematics and Science did not show stronger relationships with innovative teaching involving technology. As our sample was limited to lessons from a secondary school, our results are hardly conclusive. However, our results suggested that if creating technologically-rich learning environments for STEM subjects is a goal for innovative teaching, there are still much room for school improvement.

Inspiring teaching may emerge in the early stage of teaching when a teacher still has not shown exceptionally strong in his/her professional knowledge. Professional knowledge might hinder the adoption of innovative teaching and hamper teaching flexibility. Flexibility may be the key focus for future professional development because there is a dilemma for teachers in choosing flexibility in teaching and a better impression of teaching quality. We wish teachers to think out of the box, be flexible and be capable of reflective thinking and organise collaborative learning. Thus, we need to support them with achieving these goals without running into the risks of losing control in class.

5.3 *Limitations*

The project was small, with the number of lessons for analysis significantly reduced from the initial project plan of 300 lessons to 97 because of limited financial and human resources. Nevertheless, it was estimated that the current sample size would still be sufficient to perform the statistical analyses without sacrificing the benefit of comparing instruments developed for different purposes. This strategy was considered worthwhile and consistent with the research strategy on instrument comparison in the researcher’s previous projects. Our study is an initial step to define inspiring teaching and its outcomes, and we cannot claim that our results can resolve the problem of an overall lack of clarity and agreement completely.

5.4 *Significance*

These findings contribute to academic and professional communities in linking effective and inspiring teaching practices. The clustering results showed that teaching behaviours associated with inspiring teaching had a different pattern from effective teaching. The multiple regression results further indicated that inspiring

teaching showed a distinct group of teaching practices differing from effective teaching and impacts student engagement and the judgement of overall teaching quality differently. The CETIT seems to be a reliable tool to support researchers to study inspiring teaching in more diverse contexts, particularly in subjects like mathematics, science, language arts, art and music, where inspirations to students are found significant.

The current findings are also readily comparable with the findings in previous video studies on TIMSS lessons (e.g., Stigler et al., 1999; Seidel & Prenzel, 2006; Janik & Seidel, 2009) and a video study by the OECD on the teaching practice in nine economies (OECD, 2020). Inspiring teaching practices at secondary schools are crucial indicators of a paradigm shift in secondary education (Cheng & Mok, 2008). They also show the extent of pedagogical innovation after major curriculum reforms are introduced (Lee, 2014). Finally, the newly developed instrument will help researchers study inspiring teaching in more diverse contexts, particularly in subjects like mathematics, science, language arts, art and music, where inspirations to students are found necessary.

6 Conclusion

This study confirmed that inspiring teaching has a different pattern from that of effective teaching. The comparisons between the CETIT and ICALT indicated that the two high-inference instruments were similar in theoretical conceptualisations, administration, and reliability. While the latter looks generic, the former has a broader spectrum of teaching practices and higher relevance for observing lessons and contexts where innovative teaching, reflective thinking, flexibility and student collaboration are expected. Thus, the CETIT may have the advantage of incorporating a component associated with the inspiring teaching characteristics if a researcher has to choose only one instrument for research.

References

- Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher-student interactions in secondary school classrooms: Predicting student achievement with the classroom assessment scoring system—Secondary. *School Psychology Review*, 42(1), 76–98.
- Bacher-Hicks, A., Chin, M. J., Kane, T. J., & Staiger, D. O. (2019). An experimental evaluation of three teacher quality measures: Value-added, classroom observations, and student surveys. *Economics of Education Review*, 73, 101919.
- Cantrell, S., Fullerton, J., Kane, T. J., & Staiger, D. O. (2008). National board certification and teacher effectiveness: Evidence from a random assignment experiment (No. w14608). National Bureau of Economic Research.

- Casabianca, J. M., McCaffrey, D. F., Gitomer, D. H., Bell, C. A., Hamre, B. K., & Pianta, R. C. (2013). Effect of observation mode on measures of secondary mathematics teaching. *Educational and Psychological Measurement, 73*(5), 757–783.
- Cheng, Y. C., & Mok, M. M. (2008). What effective classroom? Towards a paradigm shift. *School Effectiveness and School Improvement, 19*(4), 365–385.
- Creemers, B., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. Routledge.
- Day, C., Sammons, P., Kington, E., & Regan, E. K. (2008). *Effective classroom practice (ECP): A mixed-method study of influences and outcomes*. ESRC.
- Ellett, C. D., & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspectives from the USA. *Journal of Personnel Evaluation in Education, 17*(1), 101–128.
- Garrett, R., & Steinberg, M. P. (2015). Examining teacher effectiveness using classroom observation scores: Evidence from the randomisation of teachers to students. *Educational Evaluation and Policy Analysis, 37*(2), 224–242.
- Gitomer, D., Bell, C., Qi, Y., McCaffrey, D., Hamre, B. K., & Pianta, R. C. (2014). The instructional challenge in improving teaching quality: Lessons from a classroom observation protocol. *Teachers College Record, 116*(6), 1–32.
- Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. National Comprehensive Center for Teacher Quality.
- Hafen, C. A., Hamre, B. K., Allen, J. P., Bell, C. A., Gitomer, D. H., & Pianta, R. C. (2015). Teaching through interactions in secondary school classrooms: Revisiting the factor structure and practical application of the classroom assessment scoring system–secondary. *The Journal of Early Adolescence, 35*(5–6), 651–668.
- Harmin, M., & Toth, M. (2006). *Inspiring active learning: A complete handbook for today's teachers*. ASCD.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Ingvarson, L., & Rowe, K. (2008). Conceptualising and evaluating teacher quality: Substantive and methodological issues. *Australian Journal of Education, 52*(1), 5–35.
- Janik, T., & Seidel, T. (Eds.). (2009). *The power of video studies in investigating teaching and learning in the classroom*. Waxmann Publishing Co.
- Kane, T. J. (2014). Do value-added estimates identify causal effects of teachers and schools? The Brown Center Chalkboard.
- Kane, T. J., & Staiger, D. O. (2008). Estimating teacher impacts on student achievement: An experimental evaluation (No. w14607). National Bureau of Economic Research.
- Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Research Paper. MET Project. Bill & Melinda Gates Foundation.
- Kane, T. J., Taylor, E. S., Tyler, J. H., & Wooten, A. L. (2011). Identifying effective classroom practices using student achievement data. *Journal of Human Resources, 46*(3), 587–613.
- Kington, A., Sammons, P., Brown, E., Regan, E., Ko, J., & Buckler, S. (2014). *Effective classroom practice*. Open University Press.
- Ko, J. Y. O. (2010). *Consistency and variation in classroom practice: a mixed-method investigation based on case studies of four EFL teachers of a disadvantaged secondary school in Hong Kong*. [Unpublished Doctoral Thesis]. The University of Nottingham.
- Ko, J., & Sammons, P. (2013). *Effective teaching: A review of research and evidence*. CfBT Education Trust.
- Ko, J., Chen, W., & Ho, M. (2015). *Teacher effectiveness and goal orientation (School report No.1)*. Hong Kong Institute of Education.
- Ko, J., Sammons, P., & Kyriakides, L. (2016). *Developing instruments for studying inspiring teachers and their teaching practice in diverse contexts. Final report*. The Education University of Hong Kong.

- Ko, J., Fong, Y., & Xie, A. (2019a, August 6). *Effective and inspiring teaching in math and science classrooms: Evidence from systematic classroom observation and implications on STEM education*. WERA Annual Meeting in Tokyo, Japan.
- Ko, J., Sammons, P., Maulana, R., Li, W., & Kyriakides, L. (2019b, April). *Identifying inspiring versus effective teaching: how do they link and differ?*. The 2019 Annual Meeting of American Educational Research Association (AERA).
- Kyriakides, L., & Creemers, B. P. (2008). Using a multidimensional approach to measure the impact of classroom-level factors upon student achievement: A study testing the validity of the dynamic model. *School Effectiveness and School Improvement*, 19(2), 183–205.
- Lee, J. C. K. (2014). Curriculum reforms in Hong Kong: Historical and changing socio-political contexts. In C. Marsh, & J. C. K. Lee (Eds.), *Asia's high performing education systems* (pp. 35–50). Routledge.
- Li, T. (2015). *Principals' instructional leadership in the context of curriculum reform in China*. [Unpublished Doctoral Thesis]. The Hong Kong Institute of Education.
- Maass, K., Cobb, P., Krainer, K., & Potari, D. (2019). Different ways to implement innovative teaching approaches at scale. *Educational Studies in Mathematics*, 102(3), 303–318.
- Maulana, R., Helms-Lorenz, M., Irnidayanti, Y., Chun, S., de Jager, T., Ko, J., & Shahzad, A. (2020, Jan). *Effective teaching behaviour across countries: Does it change over time?*. 33rd International Congress for School Effectiveness and Improvement, Marrakech, Morocco.
- Muijs, D., & Reynolds, D. (2017). *Effective teaching: Evidence and practice*. Sage.
- OECD. (2014). *Measuring innovation in education: A new perspective*. *Educational research and innovation*. OECD Publishing. <https://doi.org/10.1787/9789264215696-en>
- OECD. (2020). *Global teaching InSights: A video study of teaching*. OECD Publishing. <https://doi.org/10.1787/20d6f36b-en>
- O'Leary, M. (2020). *Classroom observation: A guide to the effective observation of teaching and learning*. Routledge.
- Opdenakker, M. C., & Van Damme, J. (2007). Do school context, student composition and school leadership affect school practice and outcomes in secondary education?. *British Educational Research Journal*, 33(2), 179–206.
- Sammons, P., Thomas, S., & Mortimore, P. (1997). *Forging links: Effective schools and effective departments*. Sage.
- Sammons, P., Kington, A., Lindorff-Vijayendran, A., & Ortega, L. (2014). *Inspiring teachers: Perspectives and practices*. CfBT Education Trust.
- Sammons, P., Lindorff, A. M., Ortega, L., & Kington, A. (2016). Inspiring teaching: Learning from exemplary practitioners. *Journal of Professional Capital and Community*, 1(2), 124–144.
- Schaffer, E. C., Nesselrodt, P. S., & Stringfield, S. (1994). The contributions of classroom observation to school effectiveness research. In D. Reynolds, B. Creemers, P. S. Nesselrodt, E. C., Shaffer, S., Stringfield, & C. Teddlie (Eds.), *Advances in school effectiveness research and practice* (pp. 133–150). Pergamon.
- Schleicher, A. (2015). Schools for 21st-century learners: Strong leaders, confident teachers, innovative approaches. In *International summit on the teaching profession*. OECD Publishing. https://www.oecd-ilibrary.org/education/schools-for-21st-century-learners_9789264231191-en
- Seidel, T., & Prenzel, M. (2006). Stability of teaching patterns in physics instruction: Findings from a video study. *Learning and Instruction*, 16(3), 228–240.
- Stigler, J. W., Gonzales, P., Kwanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS videotape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States*. NCES 99–074, U.S. Government Office.
- Taber, K. S. (2017). Reflecting the nature of science in science education. In K. S. Taber & B. Akpan (Eds.), *Science education* (pp. 21–37). Sense Publishers.
- Vieluf, S., Kaplan, D., Klieme, E., & Bayer, S. (2012). *Profiles of teaching practices and insights into innovation: Results from TALIS 2008*. OECD.

- van de Grift, W. (2007). Quality of teaching in four European countries. *Educational Research*, 49, 127–152.
- van de Grift, W. (2014). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement*, 25(3), 295–311.
- van de Grift, W., Helms-Lorenz, M., & Maulana, R. (2014). Teaching skills of student teachers: Calibration of an evaluation instrument and its value in predicting student academic engagement. *Studies in Educational Evaluation*, 43, 150–159.
- Walker, A., Hu, R., & Qian, H. (2012). Principal leadership in China: An initial review. *School Effectiveness and School Improvement*, 23(4), 369–399.
- Weisberg, D., Sexton, S., Mulhern, J., Keeling, D., Schunck, J., Palcisco, A., & Morgan, K. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness* (2nd ed.) New Teacher Project.
- Yim, O., & Ramdeen, K. T. (2015). Hierarchical cluster analysis: Comparison of three linkage measures and application to psychological data. *The Quantitative Methods for Psychology*, 11(1), 8–21.

Dr James Ko is an Associate Professor at the Department of Policy Leadership and Co-Director of the Joseph Lau Luen Hung Charitable Trust Asia Pacific Centre for Leadership and Change at the Education University of Hong Kong. Before his doctoral study, James was an EFL teacher for about 20 years and led two functional teams in a secondary school for 10 years. He is a recurrent grantee of the RGC and UGC grants and the principal investigator of 23 projects, collaborating with local academics and overseas researchers on 40 projects. He has supervised 14 doctoral students with 8 completed.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 27

Fostering Pupils' Deep Learning and Motivation in the Norwegian Context: A Study of Pupils' Perceptions of Mathematics Instruction and the Link to Their Learning Outcomes



Inger Marie Dalehefte and Esther Tamara Canninus

Abstract Recent international research has highlighted deep learning as an essential prerequisite for pupils to meet the global challenges of the future. This focus has drawn attention to Norwegian challenges, indicating that instruction leaves little room for pupils to engage intensively in tasks over time and to foster deep-learning processes. Thus, a new curriculum was implemented in the Norwegian educational system in the autumn of 2020 to emphasize deep learning throughout all content areas.

This study investigates how teachers provide learning conditions fostering learning and motivation processes to support pupils' learning during mathematics lessons. After their mathematics lesson, 144 pupils from 9 classes (grades 7–9) in seven schools in Norway completed a questionnaire. It consisted of items measuring their perception of the relevance of the content taught, the quality of the instruction given, the teacher's interest and enthusiasm, and the extent to which the instruction fulfilled their psychological needs for social relation, autonomy, and feeling competent.

On average, the pupils reported that they applied surface-level learning strategies rather than deep-level strategies in their mathematics lessons. They also lacked intrinsic motivation. To a large degree, pupils reported that they hardly recognised the content's relevance. The results support the focus on deep learning in the 2020 curriculum reform in Norway. Additionally, they reveal conditions worth investigating when aiming to foster pupils' deep learning and motivation.

Keywords Motivation · Deep learning · Mathematics · Curriculum · School-in

I. M. Dalehefte (✉) · E. T. Canninus
University of Agder, Kristiansand, Norway
e-mail: inger.m.dalehefte@uia.no

© The Author(s) 2023
R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_27

619

1 Introduction

Building on international research by authors such as Fullan et al. (2018), who pointed out that deep learning allows pupils to gain the skills necessary to tackle rapid changes in society, Norway has seen an increased interest in deep learning. The national curriculum in Norway thus far has been too extensive to stimulate and enable deep learning. In autumn 2020, the Norwegian government reduced the curriculum's content to facilitate deep learning and avoid curriculum overload (Norwegian Ministry of Education and Research, 2015). The new curriculum aims to foster pupils' abilities for broad, transferable skills and knowledge applicable to different subjects and tasks. Deep learning requires pupils to be actively engaged, reflect on their learning, and connect what is learned with what they already know (Norwegian Ministry of Education and Research, 2015). This constructivist view of learning considers learning as occurring in an active and communicative process. Although limiting the amount of content may be helpful, it is not guaranteed to lead pupils to engage in deeper learning processes or improve their learning outcomes. Investigating the communicative process in which learning occurs will illuminate how educators can support and stimulate learners to become actively involved, reflect, and connect their existing knowledge to new knowledge, thereby engaging in deep learning.

Despite widespread agreement that deep learning is appropriate for the school of the future, researchers have divergent understandings of the term 'deep learning' (Gilje et al., 2018). Fullan et al. (2018) argued for six global competencies that foster deep learning: character, citizenship, collaboration, communication, creativity, and critical thinking (p. 16). Others have criticized this framework for failing to consider a theory-based understanding of how pupils learn in a cognitive, social, and emotional way. Gilje et al. (2018) called for research on instruction that provides examples of how deep learning can be realized. Thus, this chapter considers cognitive and sociocultural views on deep learning and combines relevant theories to contribute to this perspective.

Our theoretical framework builds on Self-Determination Theory (SDT; Deci & Ryan, 1985) and Interest Theory (Krapp, 1999; Prenzel, 1995), which have focused on supportive learning conditions relevant for learning and motivational processes. Teachers impact pupils' learning by providing supportive learning conditions (Seidel, 2003). Nevertheless, the pupils must determine to which degree they use the supportive learning opportunities provided (Seidel et al., 2007). The way pupils experience the supportive learning conditions influences their motivation and learning processes. Moreover, pupils' perceptions of the classroom environment are positively related to their learning outcomes (de Jong & Westerhof, 2001; Seidel & Shavelson, 2007). In this chapter, we draw on data from 144 pupils' perceptions of the supportive learning conditions in their mathematics class and their cognitive and motivational learning outcomes. We aim to understand how educators can support and stimulate learners to engage in deep learning processes. The following research questions frame our study:

Research question 1: How do pupils perceive (a) supportive learning conditions, (b) their learning processes, and (c) their intrinsic motivation?

Research question 2: How do pupils' perceptions of supportive learning conditions impact (a) their perceived learning processes and (b) motivation?

In the following, we will describe the educational situation in southern Norway and the study's context, which builds the backdrop of our study. Next, we will briefly tap into deep learning, motivation, and supportive learning conditions before presenting the methods used and reporting on and discussing our findings.

2 The Need for a New Curriculum That Fosters Deep Learning

The Norwegian school system is free, public, and compulsory and lasts from grade 1 to grade 10. School is mandatory for all 6- to 16-year-old children. Following primary school, most pupils attend secondary school (grades 11–13). As 'one school for all' aiming at equal learning opportunities for all pupils, the Norwegian school has a diverse composition and an inclusive function. Norway prioritizes education and spends 6.4% of the gross domestic product (GDP) on educational institutions from primary to tertiary levels, which is the highest amount registered by the organisation for economic co-operation and development (OECD). Norway is also among the top three when it comes to the total expenditure on educational institutions per full-time equivalent pupil from primary to tertiary education (OECD, 2020). Socioeconomic factors play a minor role in pupils' achievement compared to many other countries, according to several large-scale assessment studies. The Norwegian government considers education to be highly important and has overseen frequent changes of curricula and school reforms throughout the years to ensure educational quality. Thus, the new curriculum initiated by the government focusing on deep learning (Norwegian Ministry of Education and Research, 2015) has achieved great attention and cost great effort in the educational system.

Southern Norway has, in some regions, special challenges related to living conditions and learning outcomes. These are characterized by, on average, a lower level of education and below average results on national standard achievement tests. (Statistics Norway, 2021; <https://www.ssb.no/>).

Our findings are based on data from a larger project (School-In, 2017–2020) funded by the Research Council of Norway (project 260,539). The project was initiated by five municipalities in southern Norway and was operationalized in cooperation with the University of Agder. The project aimed to investigate the role of local school development processes (Midtsundstad, 2019) related to inclusion in 1st- to 10th-grade schools. The project supplemented the region's focus on an inclusive learning environment, implying that children in kindergarten and schools should experience an inclusive learning environment that not only fosters children's social relatedness, but also strengthens their academic outcomes (Knutepunkt Sørlandet,

2015). In the School-In project, pupils answered questions regarding their perceptions of supportive learning conditions in their classroom and their learning processes and motivation. These questions also ask whether the pupils experienced a focus on deep learning. A meta-analysis about the effects of teaching on learning processes (Seidel & Shavelson, 2007) showed that research has more frequently investigated cognitive aspects than motivational-affective aspects. This meta-analysis also showed that domain-specific factors are most relevant for learning processes, regardless of domain, and for both cognitive and motivational-affective processes. Our study refers to the domain of mathematics instruction, which is frequently addressed within large-scale assessment and didactics studies. Thus, our findings will supplement current studies about mathematics instruction.

2.1 *Deep Learning*

Traditional theories distinguish between various learning processes (see Vermunt & Vermetten, 2004, for a review on patterns in pupil learning). While often overlapping, these theories distinguish between learning activities on different cognitive levels. For instance, Marton and Säljö (1976) focused on surface-level processing and deep-level processing. Other research has considered further aspects of learning processes; for example, Vermunt (1998) distinguished between a deep processing strategy, a stepwise processing strategy, and a concrete processing strategy. Others have broadened the perspective to include other domains. Pellegrino and Hilton (2012, as cited in Pellegrino, 2017) considered intra- and interpersonal domains alongside the cognitive domain. Research has shown that meaningful, deeper learning supports the transfer of knowledge and skills to other contexts, whereas surface knowledge and knowledge acquired through rote learning does not (Mayer, 2010). In a study with student teachers, Gordon and Debus (2002) showed that deep learning approaches positively impacted student teachers' self-efficacy. Research in higher education has been equivocal regarding whether students develop their learning approaches over time from surface to deeper approaches (see Asikainen & Gijbels, 2017).

Our research distinguishes between basic and deep elaborations based on research about teaching and learning processes in physics instruction (Seidel, 2003; Seidel et al., 2005). Basic elaborations include the core elementary topics that pupils must understand, constituting surface learning. Other forms of learning aim at deep elaborations, requiring pupils to know when, how, and why to apply the learning content. Those forms also expect pupils to reflect on how different aspects of a topic are connected, signalling deep learning.

2.2 *Intrinsic Motivation*

Motivation is a situational construct that can initiate and maintain learning processes (Prenzel, 1995; Ryan & Deci, 2017). Various theories address motivation, such as achievement goal theory (Ames, 1992) and expectancy-value theory (Wigfield & Eccles, 2000). Research has indicated that intrinsic motivation, in which the learning drive originates in the person, is essential for learning processes. We consider intrinsic motivation to be a continuum, in line with SDT (Deci & Ryan, 1985) and in relation to Interest Theory (Prenzel, 1995). On this continuum, motivation ranges from controlled motivation, in which action is driven and controlled by external rewards, to autonomous motivation, in which action and intent come from within the actor. Additionally, these theories mention amotivation, where little or no intention or action is present.

Numerous studies have shown the benefits of autonomous motivation over extrinsic or controlled motivation. Attaining extrinsic goals, such as rewards or popularity, leads to a lower degree of wellbeing than attaining intrinsic goals, such as personal growth and contributing to the community (Fryer et al., 2014; Kasser & Ryan, 1996; Unanue et al., 2014). Rump et al. (2017) observed that autonomous motivation types are negative predictors of pupils' intention to drop out of school. In a longitudinal study, Janke (2020) concluded that students in higher education who were intrinsically motivated for enrolment demonstrated a learning goal orientation. These students were also more satisfied with their choice of major. Students with extrinsic motivation for enrolment had more thoughts about dropping out and were less satisfied with their study over time (Janke, 2020). Studies have demonstrated that intrinsic motivation is related to the use of deep learning strategies (Krapp, 1999; Seidel, 2003). Thus, supportive learning conditions strengthening pupils' intrinsic motivation may also positively impact pupils' use of deep learning strategies. Questions remain about how teachers may create supportive learning conditions in their classroom to help pupils engage in deeper learning by elaborating on topics, enabling them to know when, how, and why to apply the learning content.

2.3 *Supportive Learning Conditions*

SDT and Interest Theory suggest various learning conditions to support learning and intrinsic motivation. SDT postulates that the extent to which three basic needs are fulfilled influences the degree to which intrinsic motivation is supported (Deci & Ryan, 1985). Interest Theory builds on SDT but adds a more specific focus and takes the person-object relationships into account (Krapp, 1999). An object can include a particular learning content, an abstract idea, or an action. Prenzel (1995) extended SDT with aspects from Interest Theory and related the theories to a class teaching situation. Our study builds on both perspectives. Below, we elaborate on

the supportive learning conditions proposed by SDT (Deci & Ryan, 1985) and the extended Interest Theory (Krapp, 1999; Prenzel, 1995).

2.3.1 Basic Needs – Self Determination Theory

Strengthening and supporting autonomous forms of motivation requires three basic psychological needs to be met, namely a sense of (1) autonomy, (2) competence, and (3) social relatedness (Ryan & Deci, 2017). Experiencing autonomy positively impacts learners' motivation (Tilga et al., 2020) and commitment to the learning process (Zhang et al., 2020). In class, pupils might experience **autonomy support** when provided with opportunities to make their own choices or when independent learning is supported. Experiencing that their competence is supported positively impacts learners' self-determination and motivation (Kiemer et al., 2018). Pupils experience **competence support** when they perceive their teacher trusting their skills, such as being given challenging but solvable tasks. Achieving **social relatedness** involves learners experiencing the class as a safe learning environment, characterized by unity and a friendly attitude towards each other. Higher levels of experienced social relatedness are positively related to pupils' psychological well-being, retention, and satisfaction with experiences during study (Boyd et al., 2020). Research has shown that these three psychological needs are unique and cannot be averaged into a single measure of 'satisfaction' (Van den Broeck et al., 2016). They are important for motivation, but also for learning processes (Seidel, 2003).

2.3.2 The Role of Person–Object Relationship – Interest Theory

The Interest Theory describes interest as a relation between a person and an object. It aims to explain how individuals develop from having situated to more persistent preferences for an object or activity. Prenzel (1995) argues for supplementing the SDT with elements from the Interest Theory and emphasizes that three aspects can foster the relation to an object (the content or activity) in class: (1) the relevance of the content, (2) the quality of instruction, and (3) the teacher's interest. **Relevance of content**, which helps pupils experience the content as meaningful, can be achieved by using authentic examples, content, or events that matter to the pupils. **Quality of instruction** provides structure and coherence of the content and clarifies how pupils are expected to approach a problem. The **teacher's interest** influences pupils' attitudes towards the content. A teacher showing interest in the content can ignite a spark of interest and motivation among pupils (Prenzel, 1995). These aspects have proven relevant in areas such as physics instruction (Seidel et al., 2007) and vocational education (Prenzel et al., 2002).

3 Method

3.1 Sample and Design

The data were collected as part of the School-In project, which ran from 2017 to 2020 (funded by the Research Council of Norway, project 260539) and focused on inclusion in a systemic manner. The Norwegian Centre for Research Data, which protects the privacy and rights of potential research participants, granted us permission to conduct our study. Participation in this study was voluntary and anonymous. The project was designed as a mixed methods study with an intervention in seven rural schools. For this chapter, we use data from the quantitative questionnaire, which was distributed before the intervention. One school was visited per semester (see Table 27.1). In total, 144 pupils responded to the questionnaire directly after their mathematics lesson. Pupils' ages ranged from 12 to 15 years ($M = 12.96$; $SD = .84$), with 48.6% being male, 47.2% being female, and 4.2% of the pupils not indicating their gender. The classes varied in size from 5 to 37 pupils (see Table 27.1).

3.2 Data Collection

To ensure we used high-quality analytical tools, we adapted items and scales from the IPN-Video Study in Physics instruction (Seidel et al., 2005). In total, we used 32 items. These items asked pupils about their *perception of supportive learning conditions* (Deci & Ryan, 1985; Prenzel, 1995), which consist of the three basic needs from SDT (Deci & Ryan, 1985): autonomy (4 items), competence (3 items), and social relatedness (5 items), as well as additional concepts from Interest Theory (Prenzel, 1995): relevance of content (3 items), quality of instruction (3 items), and teacher's interest (3 items). The items also asked pupils about *their perceived learning outcome* (Seidel, 2003) during the lesson: the extent to which they experienced basic elaborations (4 items), deep elaborations (4 items), and intrinsic motivation (3 items). The scales were translated into Norwegian and reformulated for the context

Table 27.1 Distribution of the sample across the schools

School	Class level			Total
	Grade 7	Grade 8	Grade 9	
School 1 (2017, Spring)	0	10	5	15
School 2 (2017, Autumn)	0	0	23	23
School 3 (2018, Spring)	0	8	0	8
School 4 (2018, Autumn)	0	9	0	9
School 5 (2019, Spring)	37	0	0	37
School 6 (2019, Autumn)	0	19	14	33
School 7 (2020, Spring)	19	0	0	19
Total	56	46	42	144

and purpose of this study. Pupils replied on a 6-point Likert scale ranging from 0 (*totally disagree*) to 5 (*totally agree*). Table 27.2 offers a description of the questionnaire's scales with item examples. All scales are internally consistent with Cronbach's alpha values ranging from .70 for teacher's interest to .95 for intrinsic motivation. The School-In project technical report offers complete documentation of the scales and items (Dalehefte & Midtsundstad, 2022).

3.3 Analysis

To answer our first research question, we calculated the descriptive values for each scale. We conducted a hierarchical linear regression analysis with two models to answer the second research question investigating the impact of learning conditions on pupils' learning outcomes. The first model shows the impact by considering the basic needs from SDT. The second model shows the added value of considering additional scales related to Interest Theory.

Table 27.2 Descriptives of the questionnaire's scales, including item examples

Scale	Number of items	Item example <i>Intro: During the lesson ...</i>	α	Mean	<i>SD</i>
<i>Learning conditions</i>					
Autonomy	4	... I had the opportunity to make my own choices.	.805	4.15	.88
Competence support	3	... the teacher gave trust that we were able to complete the tasks.	.775	4.15	.82
Social relatedness	5	... we had a good atmosphere in the class.	.791	4.05	.89
Relevance of content	3	... it became apparent that the learning content was important for us pupils.	.714	2.78	1.43
Quality of instruction	3	... I was informed what goals we should reach in the lesson.	.753	3.77	1.06
Teacher's interest	3	... I had the impression that the teacher thought the topic was interesting.	.699	4.08	.92
<i>Learning outcomes</i>					
Basic elaboration	4	... the essential points were evident to me.	.821	4.08	.92
Deep elaboration	4	... I thought about how different things are connected to each other.	.811	3.43	1.11
Intrinsic motivation	3	... I wanted to work more with the topic.	.947	2.40	1.53

4 Results

The results presented below focus on pupils' perceptions of their learning outcomes, particularly the extent to which they engaged in basic or deep elaborations and felt intrinsically motivated. Additionally, we present the extent to which the pupils experienced supportive learning conditions in their class. Lastly, we present our findings on the degree to which these supportive conditions have a predictive impact on the pupils' learning outcomes.

4.1 *Pupils' Perceptions of Elaboration and Supportive Learning Conditions*

The pupils in our sample reported experiencing basic elaborations to great a degree ($M = 4.08$; $SD = .92$; see Table 27.2), but they experienced deep elaborations during their lesson only to a slight extent ($M = 3.43$; $SD = 1.11$). The pupils slightly disagreed with having experienced intrinsic motivation during their lesson ($M = 2.40$; $SD = 1.53$).

The pupils experienced supportive learning conditions related to basic needs to a high degree (see Table 27.2). They experienced autonomy ($M = 4.15$; $SD = .88$) and competence support ($M = 4.15$; $SD = .82$) to a similar degree, closely followed by social relatedness ($M = 4.05$; $SD = .89$). Pupils also perceived their teacher to be interested ($M = 4.08$; $SD = .92$), but the average for instructional quality was lower ($M = 3.77$; $SD = 1.06$). The perceived relevance of the learning content ($M = 2.78$; $SD = 1.43$) showed the lowest mean value, indicating that pupils did not experience that the lesson was relevant to them.

Altogether, the pupils' responses showed that they mainly experienced basic elaboration but little deep elaboration and little intrinsic motivation. While their basic needs were fulfilled and they perceived their teacher as being interested, they perceived to a lesser degree the other conditions related to instruction (i.e., instructional quality and relevance of content).

4.2 *Predictive Value of Supportive Learning Conditions on Pupil Outcomes*

First, when examining predictors for the dependent variable *basic elaboration*, which refers to the most elementary learning outcomes, it becomes clear that including basic needs as predictors (Model I) allows competence support and social relatedness to predict basic elaborations. Adding conditions related to Interest Theory (Model II) considerably reduces the influence of basic needs. Of the basic needs, only competence support is a significant predictor ($\beta = .20$, $p < .10$). From Interest

Theory, quality of instruction is the single significant predictor ($\beta = .21, p < .05$). Model II predicts 23% of the variance of pupils' perception of having engaged in basic elaboration during their lesson.

Findings related to the dependent variable *deep elaboration*, which refers to the experience of perceiving deeper learning strategies, show that competence support and social relatedness predict deep elaborations in Model I ($\beta = .26, p < .01, \beta = .22, p < .01$ respectively, see Table 27.3). When considering the conditions related to Interest Theory (Model II), only content relevance has a significant impact ($\beta = .21, p < .10$) on the perception of deep elaborations. This model explains 32% of the variance in pupils' perceptions of deep elaborations.

Lastly, when considering *intrinsic motivation* as the dependent variable in Model I, competence support is the single significant predictor ($\beta = .46, p < .01$). When conditions related to Interest Theory are added to Model II, relevance of content also has a significant impact on pupils' experienced intrinsic motivation ($\beta = .39, p < .01$). In Model II, the impact of competence support is reduced to $\beta = .27 (p < .01)$. Model II explains 47% of the variance in pupils' experienced intrinsic motivation.

5 Conclusion and Discussion

Currently, Norway focuses on implementing a curriculum with a great emphasis on deep learning (Norwegian Ministry of Education and Research, 2015). Gilje et al. (2018) emphasised that various international research and trends have influenced the term *deep learning*, which has multiple meanings. Above all, Fullan et al. (2018) have influenced how the term *deep learning* is understood in Norway. Gilje et al. (2018) noted that deep learning concerns pupils' ability to develop their understanding of concepts within a subject area and be able to work in and across subject areas through problem-solving strategies and reflection. They also identified a need to understand how deep learning can be realised in instruction. In response, we applied a sociocultural perspective considering both cognitions and social interaction as essential for pupils' learning in our investigation of mathematics lessons. We studied both cognitive and motivational outcomes, as recommended by Seidel and Shavelson (2007). Thereby, we focused on supportive learning conditions based on two relevant theories about interest and motivation (Deci & Ryan, 1985; Prenzel, 1995).

First, our findings reveal that the pupils in our sample mainly experienced basic elaborations and some deep elaborations in mathematics instruction during the lesson studied. These pupils also showed little intrinsic motivation during the studied lesson. Thus, these findings are in line with the Norwegian government's recent initiatives related to the necessity of implementing a curriculum with a focus on deep learning (Norwegian Ministry of Education and Research, 2015). Second, we stated that the pupils in our sample reported perceiving supportive learning

Table 27.3 Regression coefficients of supportive learning conditions on basic elaborations, deep elaborations, and intrinsic motivation

Variable	Model I			Model II		
	B	SE B	β	B	SE B	β
<i>Basic elaborations</i>						
Constant	2.08	.42		1.86	.42	
1. Autonomy	-.02	.11	-.02	-.11	.12	-.11
2. Competence support	.32	.12	.30***	.22	.13	.20*
3. Social relatedness	.19	.09	.19 **	.12	.09	.13
4. Relevance of content				-.03	.06	-.04
5. Quality of instruction				.18	.09	.21**
6. Teacher's interest				.17	.11	.18
R ² (ΔR ²)	.17				.23 (.06)	
Adjusted R ²	.15				.20	
F	8.81***				6.41***	
<i>Deep elaborations</i>						
Constant	.28	.48		.26	.49	
1. Autonomy	.17	.13	.13	.09	.13	.07
2. Competence support	.35	.14	.26**	.18	.14	.14
3. Social relatedness	.26	.11	.22**	.16	.11	.13
4. Relevance of content				.12	.07	.15*
5. Quality of instruction				.11	.10	.10
6. Teacher's interest				.18	.12	.15
R ² (ΔR ²)	.26				.32 (.06)	
Adjusted R ²	.24				.29	
F	15.34***				10.16**	
<i>Intrinsic motivation</i>						
Constant	-2.18	.65		-1.84	.60	
1. Autonomy	.11	.17	.07	.04	.16	.03
2. Competence support	.86	.19	.46***	.50	.18	.27***
3. Social relatedness	.14	.14	.08	-.10	.13	-.06
4. Relevance of content				.41	.09	.39***
5. Quality of instruction				.11	.13	.07
6. Teacher's interest				.22	.15	.13
R ² (ΔR ²)	.31				.47 (.16)	
Adjusted R ²	.29				.44	
F	19.45***				18.96***	

Note. *N* = 144. We examined the impact of supportive learning conditions on basic elaborations, deep elaborations, and intrinsic motivation. In Model I, we entered the basic psychological needs to predict our dependent variables. In Model II, we entered content relevance, quality of instruction, and teacher's interest as predictors

p* < .10, *p* < .05, ****p* < .01

conditions related to all three basic needs (autonomy, competence, and social relatedness) and they recognized the teacher's interest during the lesson. These are

positive findings for the region, which has been working towards an inclusive learning environment for several years (Knutepunkt Sørlandet, 2015). Unfortunately, the pupils in our sample also reported perceiving less instruction quality and finding little relevance in the content of the lesson. Because these two aspects show an added value in predicting learning outcomes, as our analyses show, this finding should be treated as a cause for concern that should receive more attention in the future. Fullan et al. (2018) also emphasised the importance of content being meaningful to pupils for achieving deep learning.

This study also corroborates that both theories provide an added value in reflection about learning conditions in class. SDT (Deci & Ryan, 1985) combined with the Interest Theory elements (Krapp, 1999; Prenzel, 1995; particularly relevance of content and quality of instruction) gives valuable insight into how conditions in instruction coexist and to what degree they support pupils' intrinsic motivation and basic and deep elaborations, so that deep learning can be fostered. This theoretical background may help teachers develop their instruction towards deep learning by considering pupils' needs as well as the quality of instruction and the relevance of the content. Our results show that findings may differ depending on the use of a single theory or a combination of theories as a lens to study education. Therefore, researchers and policymakers may want to consider combining theories in their work to improve education.

Although the sample size was relatively small and restricted to mathematics instruction in grade 7 to 9, the findings provide initial insights into the potential of directing the attention towards making the content relevant to pupils within the new curriculum that aims at enhancing deep learning processes. Content relevance was a highly pertinent predictor for deep learning and intrinsic motivation in our sample. In the School-In project, which this study is a part of, we argue that linking a school's local context to instruction has great potential for both inclusive and learning processes. The local context means something to all pupils and is easy to relate to (Dalehefte & Midtsundstad, 2019). We claim that the use of examples and content from the local context has an untapped potential to improve the perception of content relevance. Further research including larger sample sizes and involving multiple regions is needed to investigate the extent to which our findings are generalizable. Other researchers have previously presented some similar findings (e.g., Frymier & Shulman, 1995; Schrod, 2013). Furthermore, although we used a well-established and well-studied instrument (Seidel et al., 2005) to collect our data, this study marked the first time this instrument was used in mathematics in a Norwegian context. Readers should be aware that, to meet the given time frame for the pupils to complete the questionnaire, we narrowed down the constructs addressed (i.e., quality of instruction was restricted to clarity and coherence) and selected a limited number of items per scale. This cost-benefit balance may have influenced this study's validity. Nevertheless, we believe the instrument is suitable and valid for this context based on our choice of items. Studies with more items per scale and a broader view on the studied constructs may investigate this claim more thoroughly.

Additional opportunities for further research lie in combining different data sources to paint a fuller picture of the situation at hand (see Kunter & Baumert,

2006). As we surveyed pupils from different schools, classes, and grades on different topics in mathematics, we could not use a mathematics test as an outcome measure to investigate the cognitive impact of the lesson because of bias in the comparisons. Additionally, pupils would be at risk of fatigue in either answering the survey or completing the mathematics test. Fortunately, as mentioned in the introduction, other research has shown that a positive relationship exists between pupils' perceptions of the classroom environment and their learning outcomes related to tests (de Jong & Westerhof, 2001; Seidel & Shavelson, 2007). Another valuable avenue for further exploration could be including teachers' perspectives (Kunter & Baumert, 2006; van der Schaaf et al., 2008).

All in all, the findings reveal that, in our sample, pupils' basic needs were met, but the pupils lacked motivation, experienced little deep learning, and struggled to see the relevance of the lesson content. The findings point into the direction of the need for a focus on deep learning in the 2020 curriculum reform in Norway. Additionally, they reveal conditions worth taking a closer look at when aiming to foster pupils' deep learning and motivation in class.

References

- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*(3), 261–271.
- Asikainen, H., & Gijbels, D. (2017). Do students develop towards more deep approaches to learning during studies? A systematic review on the development of students' deep and surface approaches to learning in higher education. *Educational Psychology Review, 29*(2), 205–234.
- Boyd, N. M., Liu, X., & Horissian, K. (2020). Impact of community experiences on student retention perceptions and satisfaction in higher education. *Journal of College Student Retention: Research, Theory & Practice*. <https://doi.org/10.1177/1521025120916433>
- Dalehefte, I. M., & Midtsundstad, J. H. (2019). Linking school's local context to instruction: An important characteristic of the in-service teacher professionalisation in school-in. In I. M. Dalehefte & S. Zehetmeier (Eds.), *Supporting teachers: Improving instruction. Examples of research-based teacher education* (pp. 77–88). Waxmann.
- Dalehefte, I. M., & Midtsundstad, J. H. (2022). *A systemic approach to school development – Technical report on the project school-in*. Waxmann. <https://www.waxmann.com/index.php?eID=download&buchnr=4502>
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum Press.
- De Jong, R., & Westerhof, K. J. (2001). The quality of student ratings of teacher behaviour. *Learning Environments Research, 4*, 51–85.
- Fryer, L. K., Ginns, P., & Walker, R. (2014). Between students' instrumental goals and how they learn: Goal content is the gap to mind. *British Journal of Educational Psychology, 84*(4), 612–630.
- Frymier, A. B., & Shulman, G. M. (1995). "What's in it for me?": Increasing content relevance to enhance students' motivation. *Communication Education, 44*(1), 40–50.
- Fullan, M., Quinn, J., & McEachen, J. (2018). *Deep learning – Engage the world – Change the world*. Corwin Press.

- Gilje, Ø., Flygt Landfald, Ø., & Ludvigsen, S. (2018). Dybdeløring -historisk bakgrunn og teoretiske tilnærminger [Deep learning – Historical background and theoretical approaches]. *Bedre Skole*, 4(30), 22–27.
- Gordon, C., & Debus, R. (2002). Developing deep learning approaches and personal teaching efficacy within a preservice teacher education context. *British Journal of Educational Psychology*, 72(4), 483–511.
- Janke, S. (2020). Prospective effects of motivation for enrollment on Well-being and motivation at university. *Studies in Higher Education*, 45(12), 2413–2425.
- Kasser, T., & Ryan, R. M. (1996). Further examining the American dream: Differential correlates of intrinsic and extrinsic goals. *Personality and Social Psychology Bulletin*, 22(3), 280–287.
- Kiemer, K., Gröschner, A., Kunter, M., & Seidel, T. (2018). Instructional and motivational classroom discourse and their relationship with teacher autonomy and competence support—Findings from teacher professional development. *European Journal of Psychology of Education*, 33(2), 377–402.
- Knutepunkt Sørlandet. (2015). *Rammeverk for satsingen 'Inkluderende Læringsmiljø'* [Framework for the initiative 'Inclusive Learning Environment']. Knutepunkt Sørlandet.
- Krapp, A. (1999). Interest, motivation and learning: An educational-psychological perspective. *European Journal of Psychology of Education*, 14(1), 23–40.
- Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, 9(3), 231–251.
- Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I-outcome and process. *British Journal of Educational Psychology*, 46(1), 4–11.
- Mayer, R. (2010). Applying the science of learning to instruction in school subjects. In R. J. Marzano (Ed.), *On excellence in teaching* (pp. 93–111). Solution Tree Press.
- Midsundstad, J. H. (2019). *Lokal skoleutvikling. Sammenhengen mellom sted, roller og undervisning* [Local school development. The connection between place, roles and instruction]. Fagbokforlaget.
- Norwegian Ministry of Education and Research. (2015). *The school of the future – Renewal of subjects and competences*. Official Norwegian Report (NOU 2015:8). <https://www.regjeringen.no/contentassets/da148fec8c4a4ab88daa8b677a700292/en-gb/pdfs/nou201520150008000engpdfs.pdf>
- Organisation for economic co-operation and development. (2020). *Education at a glance 2020: OECD-indicators*. OECD Publishing.
- Pellegrino, J. (2017). Teaching, learning and assessing 21st century skills. In S. Guerriero (Ed.), *Pedagogical knowledge and the changing nature of the teaching profession* (pp. 223–251). OECD Publishing. <https://doi.org/10.1787/9789264270695-12-en>
- Prenzel, M. (1995). Zum Lernen bewegen. Unterstützung von Lernmotivation durch Lehre [Movement towards learning. Supporting learning motivation through instruction]. *Blick in die Wissenschaft*, 4(7), 58–66.
- Prenzel, M., Kramer, K., & Drechsel, B. (2002). Self-determined and interested learning in vocational education. In K. Beck (Ed.), *Teaching-learning processes in vocational education: Foundations of modern training programmes* (pp. 43–68). Peter Lang.
- Rump, M., Esdar, W., & Wild, E. (2017). Individual differences in the effects of academic motivation on higher education students' intention to drop out. *European Journal of Higher Education*, 7(4), 341–355.
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Publications.
- Schrodt, P. (2013). Content relevance and students' comfort with disclosure as moderators of instructor disclosures and credibility in the college classroom. *Communication Education*, 62(4), 352–375.
- Seidel, T. (2003). *Lehr-Lernskripts im Unterricht* [Teaching and learning scripts in instruction]. Waxmann.

- Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77*(4), 454–499.
- Seidel, T., Prenzel, M., & Kobarg, M. (2005). *How to run a video study*. Waxmann.
- Seidel, T., Prenzel, M., Rimmele, R., Herweg, C., Kobarg, M., Schwindt, K., & Dalehefte, I. M. (2007). Science teaching and learning in German physics classrooms—findings from the IPN-video study. In M. Prenzel (Ed.), *Studies on the educational quality of schools. The final report on the DFG priority Programme* (pp. 79–99). Waxmann.
- Statistics Norway. (2021). *Tabell 08921: Utdanningsnivå etter fylke, alder og kjønn* [Table 08921: Educational level by county, age, and gender]. <https://www.ssb.no/statbank/table/08921/>
- Tilga, H., Kalajas-Tilga, H., Hein, V., Raudsepp, L., & Koka, A. (2020). How teachers' controlling behaviour can ruin students' intrinsic motivation in a physical education lesson: Test of a conditional process model. *International Journal of Sport Psychology, 51*, 81–99.
- Unanue, W., Dittmar, H., Vignoles, V. L., & Vansteenkiste, M. (2014). Materialism and Well-being in the UK and Chile: Basic need satisfaction and basic need frustration as underlying psychological processes. *European Journal of Personality, 28*(6), 569–585.
- Van den Broeck, A., Ferris, D. L., Chang, C.-H., & Rosen, C. C. (2016). A review of self-determination theory's basic psychological needs at work. *Journal of Management, 42*(5), 1195–1229.
- Van der Schaaf, M. F., Stokking, K. M., & Verloop, N. (2008). Teacher beliefs and teacher behaviour in portfolio assessment. *Teaching and Teacher Education, 24*, 1691–1704.
- Vermunt, J. D. (1998). The regulation of constructive learning processes. *British Journal of Educational Psychology, 68*(2), 149–171.
- Vermunt, J. D., & Vermetten, Y. J. (2004). Patterns in student learning: Relationships between learning strategies, conceptions of learning, and learning orientations. *Educational Psychology Review, 16*(4), 359–384.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81.
- Zhang, D., Bobis, J., Wu, X., & Cui, Y. (2020). The effects of an autonomy-supportive teaching intervention on Chinese physics students and their teacher. *Research in Science Education, 50*(2), 645–671.

Inger Marie Dalehefte is an associate professor at the University of Agder in Kristiansand/Grimstad (Norway). Her previous work at the Leibniz Institute for Science and Mathematics Education (IPN) in Kiel (Germany) mainly concerned research from the IPN Video Study in physics instruction and the professional teacher development program SINUS for Primary School with a focus on mathematics and science. Her present work addresses the school-development program School-In at the University of Agder. Her main areas of interest in research are improving instruction, professional development, educational leadership, and assessment and evaluation within the field of education. Email: inger.m.dalehefte@uia.no

Esther Tamara Canrinus is a professor at the Department of Education, University of Agder, Norway. She previously worked at the Knowledge Center of Education as a part of the Research Council of Norway, where she collaborated on writing review studies commissioned by the Norwegian Government. She also worked as a researcher and teacher educator at the University of Groningen in the Netherlands. Her research focuses on the coherence and quality of teacher education, teachers' professional development, and their professional identity. She is, furthermore, interested in teachers' social networks, classroom behaviour, and teachers' and students' motivation.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 28

The Influence of Science Teachers' Beliefs and Practices on Students' Learning Spaces and Processes: Insights from Singapore



Yuen Sze Michelle Tan and Imelda Santos Caleon

Abstract Implicit within the reform efforts in Science Education is the necessity for teachers to shift from transmissionist approaches to constructivist teaching approaches; the former emphasizes unproblematic transfer of a fixed set of ideas from credible sources to students while the latter puts primacy on students' role in knowledge construction. Teachers' beliefs may influence the implementation of reform initiatives; conversely, enactment of reform efforts may affect teachers' beliefs. Teachers' beliefs about teaching and learning, and their perceptions of their students, have been the subject of a limited and yet expanding body of research that intends to enhance the likelihood of enacting curriculum reforms that can promote students' meaningful learning. The focus of this article is to understand how teachers' beliefs influence classroom decisions that determine students' learning spaces and processes within the context of implementing school reforms.

Keywords Teacher beliefs · Science education · Teacher practices · Constructivist learning · Academic tracking

Y. S. M. Tan (✉)

Department of Curriculum and Pedagogy, University of British Columbia,
Vancouver, BC, Canada
e-mail: michelle.tan@ubc.ca

I. S. Caleon

National Institute of Education, Nanyang Technological University, Singapore,
Nanyang Ave, Singapore
e-mail: imelda.caleon@nie.edu.sg

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_28

635

1 The Influence of Science Teachers' Beliefs and Practices on Students' Learning Spaces and Processes: Insights from Singapore

New knowledge and experiences associated with curriculum reforms are often interpreted through teachers' beliefs concerning learners, classrooms, and teaching/learning materials (Pajares, 1992). Studies have shown that teachers' beliefs are useful indicators and powerful filters to help direct teachers' decisions and classroom practices (Belo et al., 2014), and are the determinants of the success of reform initiatives (Bryan, 2012; Yerrick et al., 1997). Our study aims to extend current knowledge on teachers' beliefs and practices by situating it in a dynamic education system that requires teachers to constantly adapt to new initiatives, teaching practices and assessment methods. The system follows an achievement-based process of placing incoming secondary students to different academic courses, which adopt different curricula and national assessments. Understanding teachers' beliefs and teaching practices within this less chartered research terrain may yield novel insights and surface concerns for both researchers and practitioners.

The overarching research question for the study is: What are in-service Physics teachers' beliefs and pedagogical practices of teaching the topic of electricity in the context of Singapore secondary schools?

2 Theoretical Framework and Review

Teachers' beliefs about teaching and learning, and their perceptions of their students, have been the subject of a limited and yet expanding body of research that intends to enhance the likelihood of enacting curriculum reforms that promote students' meaningful learning. Teachers' beliefs about students' learning may be categorized in accordance with the constructivist learning paradigm or, its counterpart, the absorptionist learning paradigm. The constructivist learning paradigm, which underpins current reform initiatives in science education, posits that knowledge is constructed by learners through their own conscious and personal efforts; that is, learners need to play an active, rather than passive role for meaningful learning to take place (Kruckeberg, 2006). Viewing learners as active participants in their learning, teachers provide opportunities for students to actively engage in science activities and to increase ownership in what is being learned (Kang & Wallace, 2004). Teachers create environments that are conducive for students' exploration, dialogues (Yerrick et al., 1997), and exposure to problem-solving, critical thinking and scientific argumentation (McNeill et al., 2016).

In accordance with the absorptionist learning perspective, teachers may perceive learning as a passive activity whereby learners receive knowledge from sources such as textbooks or teachers. Learners are perceived as mere recipients of knowledge and having minimal contribution to the knowledge production (see also Zohar,

2008). The transfer of knowledge from source to learners is viewed as unproblematic – knowledge is regarded as a fixed package that can be delivered to the learner unchanged (Mansour, 2013; Yerrick et al., 1997). Teachers who generally adopt this view of learning tend to emphasize students listening and taking down notes when the teachers present the ideas to be learnt.

Teachers' beliefs about their students may influence their translation of reform efforts into classroom instruction (Bryan, 2012). Teachers who are regarded as exhibiting pedagogical sensitivity take into consideration students' characteristics along with other school-related factors in making instructional decisions (Belo et al., 2014). However, teachers' perceptions of students' abilities may also limit the amount and type of reform-based practices that are enacted in the classroom (Prawat, 1992). For example, teachers who believe that their students are not capable of solving problems on their own tend to implement less inquiry activities (Lotter et al., 2007). Teachers' beliefs concerning the need to maintain the rigor of the curriculum (Kang & Wallace, 2004; Lotter et al., 2007) may serve as an obstacle to actualize curricular reforms. Teachers' motivation to adopt reform-based pedagogical approaches can be negatively affected by the pressure of having to cover content, as well as the need to strike a balance between an obligation to the discipline and to the learners when designing instructional activities (Munby et al., 2000; Yerrick et al., 1997).

3 Method

3.1 *Context of the Study: Science Education in Singapore*

This study is situated in the educational landscape of Singapore, where educational reform efforts are constantly introduced to improve the quality of education. One of the key features of the Singapore educational system is the placement of students into three academic courses – Express, Normal Academic (N[A]) and Normal Technical (N[T]) – based on the aggregate scores obtained at the Primary School Leaving Examination (PSLE). The PSLE is a national examination given at the end of elementary education (Ministry of Education [MOE], 2021).¹ Students placed in the Express course have higher aggregate PSLE scores (MOE, 2021) and, thus, are frequently perceived as academically stronger than those who qualified for the other courses. The main aims of this placement model in Singapore are to cater to individual strengths and interests of students (MOE, 2021), to help teachers cope with the diverse abilities of students, and to enable students to progress at their own pace

¹The indicative range of aggregate scores was 188 and above for the Express, 152 to 199 for N(A), and 159 and below for N(T) course in 2019 (MOE, 2021). The cut off scores for each stream may vary slightly across schools and school year.

Table 28.1 Science Curricula and National Practical Examinations for the Different Academic Courses^a

	Express “Pure science” course (Physics, Chemistry, and Biology subjects offered)	Express “Combined science” course (Combination of science subjects offered)	Normal academic science course (Combination of science subjects offered)
Curricular content	Same topics but greater depth and range of content, 4 years to cover content	Same topics but reduced depth and range of content, 4 years to cover content	Same topics but reduced depth and range of content, 5 years to cover content
National Practical Examinations (e.g., for the physics syllabus)	Demonstrate abilities to follow procedures of data recording, analysis and interpretation; plan <i>but not</i> execute scientific investigations; evaluate procedures	Demonstrate abilities to follow procedures of data recording, analysis and interpretation; plan <i>but not</i> execute scientific investigations; evaluate procedures	Not required to take the practical examinations in the fourth year

^a See UCLES-IE (2012a, b)

(Ong & Dimmock, 2013).² The curricula and assessments are differentiated according to the academic courses (see Table 28.1). Express students take the General Certificate of Education (GCE) Ordinary Level (O-Level) examination while Students in the N(A) or N(T) course take the GCE Normal Level (N-Level) examination suited for their course at the end of Grade 10 (MOE, 2021). Students from the N(A) course who did well in the N-Level examination can opt to go to the next grade and take the GCE Ordinary Level examination at the end of the next school year (MOE, 2021). (Please see Ong & Dimmock, 2013 for details on the potential effects of the examination-based placement model on students and teachers.)

3.2 Participants

Twelve Physics in-service teachers teaching either students in the Express-Combined Science Course (Luke, Simon, Yin, Ben, Fred, Tim, Wilda, Winnie)³ or N(A) Course (Laura, Lucy, Sunny, Zac)² consented to participate in the study. The teachers, who were between 20–49 years of age, taught in four Singapore public secondary schools and have a diverse range of teaching experiences (five teachers had less than 3 years of teaching experiences and the rest with 6 years or more.) All the teachers have at least a bachelor’s degree, have completed a 2-year teacher

²There are current efforts to infuse more flexibility in the placement of students into academic courses: for example, students who met eligibility requirements can transfer between courses or are offered certain subjects at a higher level via subject-based banding (MOE, 2021).

³These are pseudonyms.

education program, and have attended professional development programs focusing on reform-oriented instruction (i.e., inquiry-based teaching).

For each participating teacher, one class that he or she was teaching also participated in the study. There were 161 Express students and 102 N(A) students who participated in the study. The average class size for the Express classes was 20 (range was from 6 to 32) and 26 (range was 15 to 41) for the N(A) classes.

3.3 Data Collection

Researchers examining teachers' beliefs suggested collecting multiple data sources, particularly those concerning teacher talk and actions (e.g., Chen, 2008; Mansour, 2013; Kagan, 1992; Laplante, 1997; Schmid, 2018). Following the lead of these researchers, we deemed that classroom observations and semi-structured interviews were pertinent data to address the research question for this study.

For the classroom observations, we observed and video-recorded each teacher's lessons ("research lessons") while the teachers were enacting classroom lessons on the topic of electricity. During the audio-video recording, a video camera was positioned by a research assistant at the back of the classroom to minimize distraction of students' attention. We made 86 lesson observations (56 h in total, about 672 five-minute lesson segments), with at least six lessons on electricity recorded per teacher. Field notes were written by the research assistant while doing the video recording. Our field notes included descriptions of the participants' instruction and student/teacher interactions (e.g., description of simulations carried out by the teachers) in each 5-min lesson segment. We focused on the teachers' teaching instruction and their interactions with students based on the assumption that teachers' beliefs can shape their practices (Pajares, 1992) and influence the way they interact with their students (Gilakjani & Sabouri, 2017; Schmid, 2018). The notes also included other aspects of the lessons including class attendance, student behavior, lesson flow and content, which were potentially useful information when constructing the teaching profiles of teachers and contextualizing the enactment of their beliefs.

We conducted semi-structured individual teacher interviews prior to and after the lesson observations. Each interview lasted about 45 min. The first set of interviews elicited the teachers' beliefs about teaching and students learning and probed for their knowledge of instructional approaches suited for their students. The second set of interviews clarified the teachers' classroom practices observed through the audio-video recordings, providing teachers an opportunity to explain how their beliefs influenced their pedagogical decisions. All interviews were audio-recorded and transcribed verbatim.

3.4 *Data Analysis*

We implemented a thematic analysis approach (Miles & Huberman, 1994; Tan & Nashon, 2015) to help characterize the teachers' beliefs and teaching practices situated in the teaching of the topic of electricity. First, we *selected, reduced and organized the data* through an interactive reading and marking of the interview transcripts as they pertained to the teachers' views of teaching and learning and to the research question. Next, we *constructed teacher profiles* by making detailed notes of what the teachers deemed students and their roles to be, how students would learn the topic of electricity, and the teachers' pedagogical strategies; marked quotes from the interview transcripts were used to construct the profiles (Sandberg, 2005). We complemented the profiles by making detailed notes of the teachers' teaching instruction for each 5-min segment (672 segments in total) and compared them to the interview transcripts and the field notes. As we were interested in the extent to which teachers enact inquiry-related activities in the classroom, we counted the number of segments in the research lessons during which such activities were enacted. We subsequently *constructed themes* by looking for recurring commonalities, relationships, overarching patterns, and/or theoretical constructs as captured through words, phrases, common sequences and meanings in the marked parts of the interview transcripts, and as supported by the rest of the data set. The constructed themes were checked against the data set and refined whenever necessary.

In order to minimize bias and to develop a collective interpretation of the data set, we met up frequently to compare individual analyses, engaged in in-depth discussions, scrutinized each other's analysis and tested concepts together (Stake, 1995). We began the analysis only after the whole study was completed to prevent premature interpretations and construction of themes during the data collection phase (Sandberg, 2005).

4 Results

4.1 *Theme 1: Teachers Maintained Tight Control Over Students' Learning Process*

This theme focuses on the general challenges that the teachers faced in teaching the topic and how they improved on the basic aspects of teaching and learning to deal with such challenges. Several teachers whom we interviewed highlighted that students constantly face difficulties in applying different electricity equations to mathematics-related problems, and in understanding abstract scientific terms (e.g., voltage, potential difference and the differences between the two). When the teachers were probed for their teaching strategies, their responses revolved around ideas of maintaining a tight control over the lessons, which manifested in the research lessons as encouraging students to listen attentively in class and giving students

explicit instructions of what to copy down (c.f. Yerrick et al., 1997; Zohar, 2008). In several instances, notes were provided by the teachers where students were required to copy down the definitions of scientific terms or the formulae.

(1) ...when I start this topic, they must listen and copy down the relevant formulae and definitions...because if they don't even catch the beginning, it'd be very hard to carry on. (Laura)

In showing simulations and demonstrations, the teachers frequently employed the 'show-and-tell' approach. For example, Lucy used an online simulation to show parts of a closed circuit and how electrons move in the circuit. Throughout this simulation sequence (which took about 5 min), she stated what was supposed to be happening in the circuit and what students were supposed to see. The students were seldom probed for their observations, their interpretations of the observations, or the connections they were making to their prior knowledge or to everyday life (c.f. Kuntze, 2012).

(2) I'm going to now measure the potential difference across this first resistor here [while showing a simulation for two resistors arranged in series]. The value now is 4.5 V. (Teacher writes on the board). Now I'm going to measure the potential difference across the second resistor. And you realise the value is also 4.5 V. (Teacher points to the voltmeter connected to the second resistor in the simulation and then writes value on the board). Now from here, (teacher points to the values written on the board), can you see that your EMF is actually equals to $V_1 + V_2$? Ok? (Lucy, Lesson 7)

Considering the data drawn from the interview transcripts and classroom observations, it appears to us that the teachers maintained tight control over the students' learning in order to cope with the challenges of teaching the topic of electricity; the challenges included their perception of students' attention span as well as concept mastery. Our interpretation is further supported by how the teachers, when prompted to describe students' key roles in learning the topic, emphasized "listening in class" and "reading the textbook so that they will be able to ask questions and clarify when they are not clear" (Simon) (c.f. Yerrick et al., 1997). The teachers also conceptualized students' role as "remember[ing] what has been taught" (Yin) and "get[ting] the right answers" (Luke) when solving mathematics-related problems. In a similar vein, Laura asserted that students "listen[ing] and copy[ing] down the relevant formulae and definitions" (Excerpt 1) is critical to them solving mathematics-related problems that were introduced later in the topic.

When the teachers' pedagogies and beliefs are located within the inquiry-based reform in Singapore, it is of interest how the teachers appear to still hold the strong belief that conceptual learning necessarily precedes student-driven activities (see also Tan & Caleon, 2016; Prawat, 1992). What seems to be manifested were the teachers' strong inclinations to fall back on authoritative views of their roles, which appeared to be consistent with the dominant mode of pedagogy that is "didactic, routinized, and teacher fronted" (Kim et al., 2013, p. 294). We have observed the common lesson flow of teachers introducing the electrical components (e.g., batteries, wires, bulb), relating the components to electricity terms (current, voltage and potential difference), and then demonstrating to students how to set up the circuit;

only in a few cases do we see students having the opportunities to set up the circuit for themselves prior to the instructional flow described above.

4.2 Theme 2: Teachers' Pedagogical Decisions Were Influenced by Students' Course Placement, National Practical Examinations and Curricular Content

The teachers' perceptions of their students' abilities, which appeared to be tied to the academic courses that the students attended, affected their (teachers') pedagogical decisions in teaching the topic of electricity. Teachers of both Express and N(A) students frequently used terms like the more "capable", "intelligent" and "stronger" to refer to students from different courses. It can be inferred that some teachers considered academically weaker students, such as those attending the N(A) course, as having lower capability to take control and ownership of their own learning (c.f., Kang & Wallace, 2004; Kim et al., 2013). These perceptions resulted in the teachers' emphasis on students needing to pay attention and copy down teacher-determined notes (see e.g., Excerpt 1). Similarly, when teachers were probed for their limited use of inquiry-based activities in their research lessons, Laura, for example, expressed that:

(3) Scientific investigations [conflated with scientific inquiry in her case] are only feasible for academically stronger students, and thus I will not use investigations in my classes for academically weaker students. (Laura)

This differentiation of students was expressed by the teachers teaching the N(A) and Express courses. We noted how one teacher from the latter group also mentioned about the difference in the "caliber" of students and differentiated his students based on the "more [or least] intelligent ones" (Ben).

Concerning the practical application of scientific concepts and the use of mathematical formulae to solve physics problems, we noted in the transcripts and the research lessons how the teachers took on the responsibility to tell the students how to apply the concepts being taught (as highlighted in Theme 1). The teachers also demonstrated the 'correct' connections by *showing* students how to solve the problems. When probed for the reasons on why the teachers would make the connections for their students, teachers of N(A) students commonly held the perception that their students lacked the academic capacity, often mentioning how "the students cannot do it themselves" (Zac) or are "unable to see the connections" (Lucy, see Excerpt 1). Consequently, the teachers used perceptions of the students' abilities to justify their choice of pedagogical strategies – primarily the 'show-and-tell' approach.

Based on the above findings, it appears that the teachers might risk limiting the learning opportunities provided for the perceived academically weaker students. Our concern was also raised elsewhere (Prawat, 1992; Zohar et al., 2001). In our

opinion, the teachers might have interpreted 'differentiation'⁴ as analogous to implementing instruction based on their perception of students' capabilities and in accordance to the academic course in which they are placed. This view of teaching and learning is highly restrictive as students' capacity of growth is often overshadowed by a predetermined view of what they can or should be learning. As a case in point, our findings further suggest that the teachers' beliefs (such as role perception) and teaching practices were influenced by the practical assessment (that is, the National Practical Assessment; see Table 28.1), where helping students to prepare for the assessment may override their views of good science learning. For example, in perceiving his role as helping students to "score well in the assessment", Tim elaborated how carrying out scientific investigations in the students' reduced syllabus is "not so much an investigation but carrying out instructions of the experiments". What Tim meant was that confirmatory tests were emphasized, and this led him to avoid implementing practical activities that require students to plan and design scientific investigations, "because they don't have this type of questions in the exams". Similarly, Sunny omitted the design of scientific investigation from his electricity-based research lessons because it was "deviating from the normal question-answers".

Our classroom observations revealed that the teachers tended to provide the Express students more opportunities to work with science practical activities (11.5% out of 412 five-minute lesson segments) than their counterparts in the N(A) course (5.2% out of 260 five-minute classroom segments). Overall, our analysis suggests that teachers teaching Express and N(A) classes were utilizing scientific investigations as supplementary activities that were disconnected to their main classroom instruction, instead of using these activities extensively to teach the practices of science and to engage students with the acquisition of scientific knowledge (c.f., Wallace & Kang, 2004). Furthermore, the N(A) teachers tended to leave out scientific investigations from their lessons, noting that practical assessments are excluded from the N(A) curriculum. While Express students are required to (at least) design their investigations and, for some of them, to demonstrate their ability to 'properly' carry out the investigations, what has been suggested is that this might not necessarily translate to the larger vision of extensively engaging students in scientific inquiry.

4.3 Theme 3: Teachers' Awareness of Inconsistencies and Adoption of Flexible Pedagogy

Another theme emerging from the data is the teachers expressing their awareness of the inconsistencies between their actual classroom practices and *ideal* pedagogical scenarios that are consistent with science reform visions. The rationale for this

⁴This is a common term used amongst local teachers, such as those involved in the present study, to mean catering to differences in students' abilities when teaching.

deviation was articulated by some teachers as a practical way to deal with classroom realities and constraints. For instance, Wilda, who in Excerpt 4 underscored the need to carry out self-discovery activities in her classes, also alluded to how the demands of the national examinations propels her to be “more realistic” in planning her lessons and devote more time to preparing students for the assessments (Excerpt 5).

(4) I think they need to discover and you know, have the epiphany themselves... self-discovery... Let them try some simple things on their own, like in the lab or something. Then I reinforce it with, you know, the theory. And then they do some normal, simple problem solving, calculations. And I ask some questions.

(5) You really become less idealistic because you come up with the idea that students should be...on the path of self-discovery. But then later you learn that...you need to be more realistic. You need to make sure they are able to solve that 80% of the curriculum... They need to perform during the national exams. (Wilda)

Similarly, Yin described how “Ideally, we should have the investigation [inquiry-based investigations] for all [students], but due to time constraints, we fall back on ‘chalk and talk’ [style of teaching]”.

Some teachers, however, tended to demonstrate greater nimbleness when it comes to navigating their ideal and actual realities in teaching. For example, Zac expressed his intentions to adopt a flexible teaching approach to promote conceptual learning and problem-solving skills among his N(A) students. When probed for what he meant by flexible teaching approach, Zac underscored the keeping of learning opportunities open for his N(A) students, which, in the topic of electricity, would manifest as his deliberate inclusion of questions that he regarded as fitting for the academic ability of students from the Express course (“‘O’ level type of question”, “Pure Physics one”). As Zac described his pedagogy, he clearly articulated how the end-goal of his scaffolding strategy was for students to have opportunities to engage with questions of greater complexity and requiring greater analysis (“Pure Physics one”):

(6) I'll give them [students] a basic N(A) level problem [mathematics-related electricity problem as would be assessed in the N(A) national examinations]... Then next one will be medium level challenge. Then after that I increase to an O-Level type of question [typically used for assessments of Express students], and then if I feel like this class is ready for it...to give you [students] a Pure Physics one [typically used for assessments of high ability Express students]. (Zac)

Within an educational system that utilizes achievement-based placements as a means to cater to students’ diverse needs and abilities, Zac’s efforts suggest the feasibility of employing an instructional approach that expands (rather than limits) students’ learning spaces. It is however noteworthy that despite Zac’s efforts to provide a wide range of learning opportunities for N(A) students, his research lessons were observed to be heavily didactic. This tension draws attention to and underscores the need to be empathetic towards teachers who face challenges in reconciling their beliefs and pedagogical actions (Bryan, 2012). Within the Singaporean educational context, it also supports previous studies that highlighted

how Singapore science teachers tend to prefer more authoritative, teacher-centred styles when developing their students' scientific knowledge (Tan & Hong, 2014; Yeo & Tan, 2010). What is encouraging is that Zac was able to set a good starting point that he and other teachers may follow through with more efforts to deepen their pedagogical awareness and increase their repertoire of pedagogical activities, in order to better cater to their students' learning needs.

Phrased differently, the teachers' espoused pedagogical intentions could, on one hand, reveal a perceived gap between the 'practical instruction' and the 'ideal instruction'; this could be indicative of the misalignment between the nation's educational priorities of engaging student in scientific inquiry and the actual practices that are deeply embedded within the content- and assessment-driven nature of the educational system. On the other hand, it also highlights the ways by which teachers are adapting to this nuanced educational setting. Indeed, we share the empathetic view of Lee (2008) arguing that Singapore Science teachers in his study have enacted teaching through '*in-between spaces*' (Lee, 2008, p. 931): between policies and their own classroom teaching to infuse science learning in ways about which they are passionate. The juxtaposition of teachers' ideal views, realities and constraints can be a step for teachers towards exerting their agentic control (Brandt, 2012) that best utilizes mandated curriculum to complement their teaching and learning goals.

5 Conclusion and Discussion

In this paper, some interesting insights were drawn from the examination of teachers' beliefs and classroom actions. Although we limited our exploration to the topic of electricity, the teachers have often responded to our interview questions by describing their practices and beliefs in more general terms, that is, to describe their overall teaching. This enabled us to draw implications both for teaching of the topic and beyond, although we could definitely benefit from more studies to increase the generalizability of the results.

5.1 *Considerations for Teaching and Learning*

The nuanced understandings that emerged from the study are helpful to further unpack the impact of policies in a tightly coupled system where stipulated curriculum and national examinations are known to have profound influences on teaching and learning. We learned from the study that the participating teachers' perceptions of students' academic abilities, which were made more explicit by the placement of students to different courses characterized by differentiated curricula, could cause tensions in the ways by which teachers make their pedagogical choices. This tension is similarly reported in Wallace and Kang's (2004) study where the teachers held

two competing sets of beliefs: beliefs constraining inquiry-based teaching were more public and culturally based (including policy-based decisions), while those that promoted inquiry were more private and based on teachers' ideas of successful science learning.

Similar to Wallace and Kang (2004), we assert for the need to help teachers resolve this tension. We see glimpses of the teachers' creative agentic control as they reconcile their own learning goals for their students with mandated ones. As a case in point, we observed how teachers in our study tended towards maintaining tight classroom control. On one hand, this may be interpreted as teachers holding views of the 'old dichotomy' (Prawat, 1992), positioning themselves as the key source of knowledge, emphasizing the use of curriculum resources and/or would attempt to deal with difficult aspects of teaching the topic by ensuring that students learn key content and concepts. Within this framing, the inclusion of reform-based (inquiry) skills in national assessment may also be interpreted as being strategic but inadequate to support reform efforts, and thus warrant greater attention. On the other hand, the teachers' pedagogical decisions could be framed as an artifact of the cultural factors that guide classroom practices (Bryan, 2012), where authoritative figures such as teachers are highly respected in Asian cultures. Within this vein, we align our findings with earlier works, such as Tan and Hong's (2014), which explained the tight classroom control, a dominant form of classroom teaching in Singapore, as "a tight framing of knowledge" (p. 689) to ensure that scientific knowledge is accurately presented (see also Yeo & Tan, 2010). This could, in turn, be deemed as stemming from the teachers' private beliefs about good science teaching. Framed this way, the snapshot of the teachers' beliefs and teaching instruction captured in this study points to a compromise strategy the teachers employed in order to maintain students' learning spaces in spite of external challenges. We speculate that (in the context of this study) exerting tight classroom control could be a manifestation of the teachers' sensitivity towards their students' learning needs, rather than a neglect of them.

Another key tension teachers need to resolve stems from how they were very much bound to their obligations to prepare students for practical examinations, despite recent changes in the science curricula and practical examinations. In our opinion, this may not be perceived as an inadequacy on the part of the teacher, as it signals their responsiveness to the needs of their students who are educated within a system that places high currency on academic achievements. A potential area of growth for the teachers is to be aware of how to go beyond this goal and open up learning spaces for students (like what Zac did) to gain the skills being examined as well as other valuable skills that are not necessarily assessed.

Another good starting point for teachers to address the tensions they face in teaching is by articulating the gaps between their goals of teaching and actual classroom practices. In some cases (such as Wilda's), we noted the possible tension, even discontent, teachers faced as their classroom practices risk narrowing the learning possibilities for the students, especially for academically weaker students. In other cases, such as Zac's, teachers were able to work within the constraints to find ways to use learning tasks with increasing levels of difficulties. The findings thus allude

to the degrees by which different sets of beliefs (e.g., policy-based/public vs. personal) influence the teachers' pedagogical decisions, and thus determine student learning spaces and processes. This, in turn, highlights the pertinence of building teacher's capacities, which bears implications for teacher professional development and policy.

5.2 Considerations for Policymaking and Teacher Professional Development

Given the close relationship between teachers' beliefs and educational policies as is revealed through this study, the findings highlight the benefits of exploring how reform initiatives are communicated through policy documents. First, attention is drawn to the implicit messages in various policy and framework documents that could potentially be misconstrued by teachers. We recognize how the academic placement efforts were purposed to help teachers cope with diverse learning needs of students. However, the achievement-based placement model may result in teachers' misinterpretation of the model's original intentions and the unintended consequence of students having limited access to various learning experiences; this is exacerbated by the reduction of curriculum content and national practical assessment formats (and at some point, excluding the practical examinations) in academic courses. The misalignment between policy decisions and the enactment of these policies warrants greater attention to *how teachers interpret* prescribed curricula and reform-based documents (Tan & Caleon, 2016; Tan & Nashon, 2015). There is also a need for greater coherence and coordination between science curricula, assessment and instruction for both students and teachers, as was asserted by Duschl et al. (2007).

Second, if the intention of the achievement-based placement process is to provide students with varying academic abilities appropriate attention and guidance, it would be imperative to build teachers' capacity to diversify students' learning experiences, utilize the resources that students bring into the classroom, and to explore a variety of ways to actively engage students in their science learning. Our findings show that this is feasible (as exemplified by Zac) within an educational setting such as Singapore's, where professional development is highly supported and often initiated by the Ministry of Education (Somekh & Zeichner, 2009). Recent studies in Singapore have reported on the benefits of teachers collaborating and engaging with research/inquiry within the loci of their own classrooms, which included helping science teachers to meaningfully integrate their beliefs with mandated curricular goals, and to fruitfully utilize the curriculum to promote teachers' desired visions of student learning (e.g., Tan & Nashon, 2013, 2015).

6 Summary

Although the findings of the present study were drawn from data collected on a small sample of teachers, the study adds to the current literature on teachers' beliefs and practices pertaining to the implementation of science reform visions, which is sparser when located within the context of a non-Western education culture. The present study paints concrete scenarios illustrating how teachers' beliefs and practices were influenced by contextual forces, such as curricular content, national assessments, and achievement-based placement process. While such contextual forces may bring about tensions between what teachers set as ideals for teaching and learning and their responsibilities to address the needs of their students, and, in some cases, limitations in the learning experiences offered to students, we have observed indications of teachers adopting a flexible, creative and contextually nuanced pedagogy. The latter serves as a good starting point to better equip teachers in broadening students' learning spaces and to optimize learning.

Acknowledgements The authors express their gratitude to the participating teachers and schools for their support in the research study. This paper refers to data from the research project OER 08/11 ISC, funded by the Education Research Funding Programme, National Institute of Education (NIE), Nanyang Technological University, Singapore. The views expressed in this paper are solely those of the authors and do not necessarily represent the views of NIE.

References

- Belo, N. A. H., van Driel, J. H., van Veen, K., & Verloop, N. (2014). Beyond the dichotomy of teacher- versus student-focused education: A survey study on physics teachers' beliefs about the goals and pedagogy of physics education. *Teaching & Teacher Education*, *39*, 89–101. <https://doi.org/10.1016/j.tate.2013.12.008>
- Brandt, C. B. (2012). Misrecognition and science education reform. *Cultural Studies of Science Education*, *7*(3), 579–583. <https://doi.org/10.1007/s11422-012-9398-y>
- Bryan, L. A. (2012). Research on science teachers' beliefs. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education* (Vol. 1, pp. 477–495). Springer. https://doi.org/10.1007/978-1-4020-9041-7_33
- Chen, C. H. (2008). Why do teachers not practice what they believe regarding technology integration? *The Journal of Educational Research*, *102*(1), 65–75. <https://doi.org/10.3200/JOER.102.1.65-75>
- Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (2007). Taking science to school: Learning and teaching science in grades K-8. *National Academy Press.*, *22*, 1265–1266. <https://doi.org/10.1007/s11191-011-9353-4>
- Gilakjani, A. P., & Sabouri, N. B. (2017). Teachers' beliefs in english language teaching and learning: A review of the literature. *English Language Teaching*, *10*, 78–86. <https://www.ccsenet.org/journal/index.php/elt/article/view/66914>
- Kagan, D. M. (1992). Implication of research on teacher belief. *Educational Psychologist*, *27*, 65–90. https://doi.org/10.1207/s15326985ep2701_6
- Kang, N. H., & Wallace, C. S. (2004). Secondary science teachers' use of laboratory activities: Linking epistemological beliefs, goals, and practices. *Science Education*, *89*(1), 140–165. <https://doi.org/10.1002/sce.20013>

- Kim, M., Tan, A. L., & Talaue, F. T. (2013). New vision and challenges in inquiry-based curriculum change in Singapore. *International Journal of Science Education*, 35(2), 289–311. <https://doi.org/10.1080/09500693.2011.636844>
- Kruckeberg, R. (2006). A Deweyan perspective on science education: Constructivism, experience, and why we learn science. *Science & Education*, 15(1), 1–30. <https://doi.org/10.1007/s11191-004-4812-9>
- Kuntze, S. (2012). Pedagogical content beliefs: Global, content domain-related and situation-specific components. *Educational Studies in Mathematics*, 79(2), 273–292. <https://doi.org/10.1007/s10649-011-9347-9>
- Laplante, B. (1997). Teachers' beliefs and instructional strategies in science: Pushing analysis further. *Science Education*, 81, 277–294.
- Lee, Y. J. (2008). Thriving in-between the cracks: Deleuze and guerilla science teaching in Singapore. *Cultural Studies of Science Education*, 3, 917–935. <https://doi.org/10.1007/s11422-008-9116-y>
- Lotter, C., Harwood, W. S., & Bonner, J. J. (2007). The influence of core teaching conceptions on teachers' argumentation instruction in their science classrooms. *Journal of Research in Science Teaching*, 44(9), 1318–1347. <https://doi.org/10.1002/tea.20191>
- Mansour, N. (2013). Consistencies and inconsistencies between science teachers' beliefs and practices. *International Journal of Science Education*, 35(7), 1230–1275. <https://doi.org/10.1080/09500693.2012.743196>
- McNeill, K. L., Katsh-Singer, R., González-Howard, M., & Loper, S. (2016). Factors impacting teachers' argumentation instruction in their science classrooms. *International Journal of Science Education*, 38(12), 2026–2046. <https://doi.org/10.1080/09500693.2016.1221547>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis* (2nd ed.). Sage. [https://doi.org/10.1016/s1098-2140\(99\)80125-8](https://doi.org/10.1016/s1098-2140(99)80125-8)
- Ministry of Education (Singapore). (2021). *Courses and subjects for secondary schools*. <https://www.moe.gov.sg/secondary/courses>
- Munby, H., Cunningham, M., & Lock, C. (2000). School science culture: A case study of barriers to developing professional knowledge. *Science Education*, 84(2), 193–211. [https://doi.org/10.1002/\(sici\)1098-237x\(200003\)84:2<193::aid-sce4>3.0.co;2-k](https://doi.org/10.1002/(sici)1098-237x(200003)84:2<193::aid-sce4>3.0.co;2-k)
- Ong, C. H., & Dimmock, C. (2013). Principals' engagement of low ability students in Singapore secondary schools. *Educational Management Administration & Leadership*, 41(2), 214–232. <https://doi.org/10.1177/1741143212468345>
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307–332. <https://doi.org/10.3102/00346543062003307>
- Prawat, R. S. (1992). Teachers' beliefs about teaching and learning: A constructivist perspective. *American Journal of Education*, 100(3), 354–395. <https://doi.org/10.1086/444021>
- Sandberg, J. (2005). How do we justify knowledge produced within interpretive approaches? *Organizational Research Methods*, 8(1), 41–68. <https://doi.org/10.1177/1094428104272000>
- Schmid, R. (2018). Pockets of excellence: Teacher beliefs and behaviors that lead to high student achievement at low achieving schools. *SAGE Open*, 8, 215824401879723. <https://doi.org/10.1177/2158244018797238>
- Somekh, B., & Zeichner, K. (2009). Action research for educational reform: Remodelling action research theories and practices in local contexts. *Educational action research*, 17(1), 5–21. <https://doi.org/10.1080/09650790802667402>
- Stake, R. E. (1995). *The art of case study research*. Sage.
- Tan, Y. S. M., & Caleon, I. S. (2016). Dealing with the ambiguities of science inquiry. *The Physics Teacher*, 54(3), 168–171. <https://doi.org/10.1119/1.4942141>
- Tan, A. L., & Hong, H. (2014). Learning science in high school: What is actually going on? *International Journal of Science Education*, 36(4), 677–692. <https://doi.org/10.1080/09500693.2013.823676>

- Tan, Y. S. M., & Nashon, S. M. (2013). Promoting teacher learning through learning study discourse: The case of science teachers in Singapore. *Journal of Science Teacher Education*, 24(5), 859–877. <https://doi.org/10.1007/s10972-013-9340-5>
- Tan, Y. S. M., & Nashon, S. M. (2015). Promoting teachers' collaborative exploration of a new science curriculum: The case of a Singapore learning study. *Professional Development in Education*, 41(4), 671–689. <https://doi.org/10.1080/19415257.2014.944670>
- University of Cambridge Local Examinations Syndicate International Examinations (UCLES-IE). (2012a). *Examination syllabuses for 2013: UCLES-IE ordinary level subjects*. University of Cambridge.
- University of Cambridge Local Examinations Syndicate International Examinations (UCLES-IE). (2012b). *Examination syllabuses for 2013: Normal level (academic) & (technical) subjects*. University of Cambridge.
- Wallace, C. S., & Kang, N. H. (2004). An investigation of experienced secondary science teachers' beliefs about inquiry: An examination of competing belief sets. *Journal of Research in Science Teaching*, 41(9), 936–960. <https://doi.org/10.1002/tea.20032>
- Yeo, J., & Tan, S. C. (2010). Constructive use of authoritative sources in science meaning-making. *International Journal of Science Education*, 32(13), 1739–1754. <https://doi.org/10.1080/09500690903199564>
- Yerrick, R., Parke, H., & Nugent, J. (1997). Struggling to promote deeply rooted change: The 'filtering effect' of teachers' beliefs on understanding transformational views of teaching science. *Science Education*, 81(2), 137–159. [https://doi.org/10.1002/\(sici\)1098-237x\(199704\)81:2<137::aid-sce2>3.0.co;2-g](https://doi.org/10.1002/(sici)1098-237x(199704)81:2<137::aid-sce2>3.0.co;2-g)
- Zohar, A. (2008). Science teacher education and professional development in argumentation. In S. Erduran & M. P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 245–268). Springer. https://doi.org/10.1007/978-1-4020-6670-2_12
- Zohar, A., Degani, A., & Vaaknin, E. (2001). Teachers' beliefs about low-achieving students and higher order thinking. *Teaching and Teacher Education*, 17(4), 469–485. [https://doi.org/10.1016/S0742-051X\(01\)00007-5](https://doi.org/10.1016/S0742-051X(01)00007-5)

Yuen Sze Michelle Tan is an Associate Professor of Science Education in the Department of Curriculum and Pedagogy at the University of British Columbia. Her research focuses on science teacher education and she takes special interest in teachers' engagement with reform-associated pedagogies. Her research is located in different teaching and learning contexts, including different educational systems and a variety of collaborative teacher inquiry models. Her current projects include integrating educational neuroscience with classroom practices and promoting social action through culturally responsive and community-based science education. Email: michelle.tan@ubc.ca

Imelda Santos Caleon is an Assistant Dean for Partnerships and a Senior Research Scientist at the Office of Education Research, National Institute of Education, Nanyang Technological University, Singapore. Her research interests are in the areas of positive education and science education, with a particular focus on resilience, mindsets, and metacognition. Her foremost intent was to utilize and develop approaches rooted in positive psychology to facilitate learners' conceptual and mindset change, and build resources (emotional, psychological, social and cognitive) that can help learners, especially students placed at risk, to thrive in school and beyond. Her recent research projects focus on the examination and fostering of social and emotional well-being, metacognition, and adaptive stress mindsets of adolescents. Email: imelda.caleon@nie.edu.sg

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 29

The Illusion of Perspective: Examining the Dynamic Between Teachers' Perceived and Observed Effective Teaching Behaviour



Benjamin Looker, Alison Kington, Kimberley Hibbert-Mayne,
Karen Blackmore, and Scott Buckler

Abstract Effective teaching behaviour is known to be associated with positive pupil outcomes. As such, it is considered an important aspect of educational research. In this chapter, we used validated instruments to measure two types of teaching effectiveness. Teachers' perceived effective teaching behaviour was measured using the Teacher as Social Context (TASC) questionnaire and teachers' observed effective teaching behaviour were measured using the International Comparative Analysis of Learning and Teaching (ICALT) Observation Instrument. Statistical comparisons were made between these two measures and were additionally analysed through the lens of teachers' career phases. The study found that there are significant differences in teachers' perceived and observed effective teaching behaviour. Teachers' perceived effective teaching behaviour was found to remain relatively stable throughout their careers, however, their observed effectiveness was seen to change considerably. As teachers enter the middle phases of their careers, an increase in observed effectiveness was identified, followed by a decline during the later career phases. Further analysis of observed effective teaching behaviour using six ICALT domains indicates that the way a teacher facilitates a safe and stimulating learning climate and is efficiently organised plays an important role in the variation of their observed effectiveness. These results have implications for the continued professional development of trainee teachers and qualified teachers at all stages of their careers.

B. Looker (✉) · A. Kington · K. Hibbert-Mayne · K. Blackmore
University of Worcester, Worcester, UK
e-mail: k.blackmore@worc.ac.uk

S. Buckler
Holy Trinity School, Worcestershire, UK

© The Author(s) 2023
R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_29

Keywords Teacher effectiveness · Perceived effectiveness · Observed effectiveness · Career phase · Effective teaching behaviour

1 Introduction

Teacher effectiveness has a long tradition of research, with various authors discussing how an effective teacher needs to not only know and understand their students but also the problems they may encounter in their learning. Effective teachers should be able to incorporate their knowledge of students into their classroom practice while respecting and encouraging learners to raise their expectations (e.g. Brown & McIntyre, 1993; Kington et al., 2014; Ruddick et al., 1996; TLRP, 2013; Upton & Taylor, 2014; Wray et al., 2000).

This chapter presents findings from a cross-sectional analysis that explored observed measures of effective teaching behaviour alongside teachers' self-reported perceptions of their classroom effectiveness obtained using a teacher questionnaire. Focusing on the final wave of data collection, observational data were examined and compared with teachers' questionnaire responses. Analyses of observed and perceived teaching effectiveness identified variations in practice depending on the length of service (or career phase) of the teacher. In addition, analysis using radar plots suggested that teachers' effective organisational skills are a key component, acting as a limiting factor to overall teaching effectiveness.

2 Conceptual Framing

2.1 *Teacher Effectiveness*

An effective education can be defined as improving student achievement (Coe et al., 2014). It is therefore unsurprising that a considerable amount of teacher effectiveness literature focuses on the relationship between teaching and student outcomes. Varying perspectives on the purposes of schooling may affect the priority placed on the different qualities, qualifications, practices, and accomplishments of teachers (Little et al., 2010). There is some agreement that the outcomes of students should not only include new learning, but progression within social, affective and psychomotor domains (Sammons, 1999; Kyriakides & Creemers, 2012); however, to be considered trustworthy, measurements of teacher effectiveness continue to be predominantly based on the academic progress of students (Coe et al., 2014). Consequently, the last few decades have seen the identification of teaching behaviours, teaching skills and other generic features of effective classroom practice which are positively related to student academic achievement (Day et al., 2007; Coe et al., 2014; Kington et al., 2014). For example, teachers' attributes and actions have been found to be associated with variance in student academic outcomes (Muijs & Reynolds, 2011; Kyriacou, 2018). However, Day et al. (2007), and more recently

Muijs et al. (2014), do not limit the characterisation of teacher effectiveness to academic outcomes and suggest that variations in school and classroom contexts (e.g. leadership, culture, colleagues, subject area and socio-economic factors) be used for measuring teacher effectiveness differently. Though the attributes and behaviours of teachers have been firmly integrated into theoretical and empirical models of educational effectiveness for decades (e.g., Creemers & Kyriakides, 2013), they are not easily characterised (Brown et al., 2014) which has potentially contributed to the predominance of teacher effectiveness literature being based on academic outcomes.

Defining teacher attributes and identifying their impact on classroom effectiveness has been linked to perspective (Coe et al., 2014). Literature specifically exploring teachers' perceptions of their effectiveness has predominantly focused on conceptual and methodological issues pertaining to teachers' beliefs in their own capability (e.g., Henson, 2010; Klassen et al., 2009; Labone, 2004; Tschannen-Moran & Woolfolk Hoy, 2001; Wyatt, 2014) and it has been suggested that this sense of self-efficacy in the classroom is an important factor in teachers' effectiveness (Caprara et al., 2003; Caprara et al., 2006; Aloe et al., 2014). For example, the VITAE¹ project (Day et al., 2007), which tracked the effectiveness of 300 primary and secondary school teachers over 4 years, found that there was a significant relationship between a teacher's perceived effectiveness (as reported through questionnaire surveys and interviews) and 'relative' effectiveness (as measured through classroom observations and student national test scores). The study also identified that teachers' perceived effectiveness strongly reflected their overall sense of self-efficacy as a practitioner. Furthermore, their analysis identified that perceived and observed effectiveness were directly affected (to varying degrees) by length of service in the profession which, in turn, affected the way teachers viewed their effectiveness, both positively and negatively, in the classroom (Day et al., 2007).

2.2 *Teacher Career Phase*

Teachers' career phases have been categorised in a variety of ways. Super's (1957) four-stage model suggested that there are distinct phases related to the length of service. Super argued that teachers move through these phases, referred to as exploration, establishment, maintenance and disengagement, although not necessarily in a linear way. Later, Huberman's (1989) research into secondary school teachers' career development expanded on Super's non-linear model and identified that teachers experience five distinct, discontinuous career phases; namely career entry, stabilization, experimentation, conservatism and disengagement (Huberman, 1993). More recently, variations in teachers' career phases have been further refined through the VITAE project (Day et al., 2007) which developed a six-phase model based on teachers' professional lives. These phases follow certain discernible

¹Variations in Teachers' Work, Lives and Effectiveness, commissioned by the Department for Education and Skills.

Table 29.1 Summary of career phase characteristics. (Derived from Day et al., 2007)

Phase	Summary of characteristics
0–3 years	High commitment and challenge. Developing a sense of efficacy in the classroom.
4–7 years	Increasing confidence in teachers' belief in their own effectiveness.
8–15 years	Managing transitions in role and identity. High sustained engagement.
16–23 years	Work-life tensions leading to challenges in motivation and commitment.
24–30 years	Facing challenges to sustaining motivation.
31+ years	Maintaining commitment with some declining commitment.

patterns characterised by identifiable stages (Day et al., 2007) which are summarised in Table 29.1.

Using the career phase as a conceptual lens, this chapter explores variations between teachers' perceived effectiveness compared with observations of their practice in England. To this end, three broad research questions were used to guide the analysis:

1. Is there a difference between teachers' *perceived* and *observed effectiveness*?
2. How do *perceived* and *observed effectiveness* vary according to teacher career phase?
3. How can variations in *observed effectiveness* across career phases be explained?

2.3 Research Context

In England, there are five stages of education, namely early years (which includes nursery and pre-school phases), primary school, secondary school, further education (post-16 years) and higher education. This study involved teachers working in primary and secondary schools where the curriculum is further divided into 'key stages' based on child age; as such key stages 1 (age 5–6 years) and 2 (age 7–10 years) are covered by the primary stage, whilst key stages 3 (age 11–13 years) and 4 (age 14–16 years) are covered by the secondary stage. General Certificates of Secondary Education (GCSEs) are taken at the end of key stage 4. In England, the majority of state-funded primary and secondary schools are mandated to follow the National Curriculum. However, since 2010, many schools have been granted 'Academy School' or 'Free School' status which allows more flexibility over the curriculum as well as independence from the local authority with regards to teacher pay and conditions. While academies and free schools have more autonomy over curriculum decisions, all state-funded schools are subject to inspection by the Office for Standards in Education (Ofsted) who expect to find learners studying a full range of subjects by teachers who have 'good knowledge of the subject', who 'present subject matter clearly', 'use assessment well' and 'create an environment that allows the learner to focus on the learning' (Ofsted, 2021: 39–40). It is worth noting that across all state-funded schools around 1 in 5 pupils are eligible for free school meals based on their socioeconomic background.

In terms of PISA results, the UK has improved in reading, moving from 25th to 14th amongst OECD countries (OECD, 2021), with England having the highest score of the four UK nations. England is also above the average OECD scores in maths and science, showing an upward trend. England has a young teaching population, with its teachers having spent fewer years in the classroom than teachers in most other TALIS countries (OECD, 2019). The average is 13 years, which ranks 46th out of the 50 countries. Only 18% of the teaching population is over 50 years of age, compared to the OECD average of 34%. Furthermore, practitioners in England report high levels of stress, with 38% of teachers reporting 'a lot' of stress in work, compared to the OECD average of 18%. More recently, the OECD reported that the UK had the second highest attrition rate of OECD countries (OECD, 2021).

2.4 Methods and Procedures

This longitudinal study between 2015 and 2019 was conducted through observations of classroom practice, using the International Comparison of Learning and Teaching (ICALT) observation instrument (van de Grift, 2007; van de Grift, 2014), and the Teacher as a Social Context (TASC) questionnaire (Wellborn et al., 1992) which explored teachers' perceptions of their own practice. The data were collected over 4 years in schools in England, with a growing number of observations conducted each year as increasing numbers of practitioners were recruited to the study. The cross-sectional data reported in this chapter were gathered in the final year of data collection, when a total of 312 lesson observations were carried out, with each teacher observed also completing a teacher questionnaire.

2.5 Instruments

2.5.1 Effective Teaching Behaviour Observations

According to Wragg (1999), classrooms are complex environments representing an interplay of variables that affect observations. The reliability and validity of several established classroom observation instruments have been questioned by various researchers (e.g. Baker et al., 2010; Biesta, 2009; Page, 2016). Furthermore, van de Lans et al. (2016) highlight the particular issue of substantial measurement error, where a judgment of a teacher may not be indicative of their overall performance if, for example, they are working with a difficult class, are feeling ill, and so forth. It could be argued that, in contrast, systematic observation tools such as the ICALT instrument are considered as a valuable method to enable the comparison of teachers' teaching behaviours; since, in addition to using standardised terms, the instrument includes pre-determined and agreed categories describing elements of observable classroom practice.

The ICALT structured observation schedule consisted of seven domains of teacher effectiveness:

1. A safe and stimulating learning climate – 4 indicators;
2. Efficient organisation – 4 indicators;
3. Clear and structured instructions – 7 indicators;
4. Intensive and activating teaching – 7 indicators;
5. Adjusting instructions and learner processing to inter-learner differences – 4 indicators;
6. Teaching and learning strategies – 6 indicators;
7. Learner engagement – 3 indicators.²

This observation tool was piloted, and inter-rater reliability was determined for 10 lessons rated independently by paired researchers. The most appropriate indicator to assess inter-rater reliability for an instrument consisting of ordinal scales, such as the ICALT tool, is the Weighted Kappa and the inter-rater reliability score was statistically significant (mean Weighted Kappa Quadratic = 0.73), which is considered highly reliable (Bakeman & Gottman, 1997). For the main study, observations were conducted by individual researchers and completed during the lesson. Lessons were observed across a range of subjects throughout all key stages (1–4). Each lesson observed lasted between 30 and 60 min. Cronbach's alpha for the ICALT observation instrument indicated excellent reliability of the scale ($\alpha = 0.95$).

2.5.2 Teacher Questionnaire

Questionnaires, designed to be administered alongside the ICALT observation tool (Maulana et al., 2014), were distributed to teachers directly after the lesson observations had been conducted, and teachers were asked to complete the survey in relation to the observed lesson. Responses were scored on four-point Likert scales ranging from 1 (strongly disagree) to 4 (strongly agree). The questionnaire gathered data according to three areas relating to different aspects of classroom practice.

The questionnaire teacher as a social context (TASC) was used to explore teachers' perceptions of their effectiveness. The 41 items in this section directly relate to the actions and behaviours of teachers and includes items associated with social aspects of teaching and self-efficacy (e.g. Tschannen-Moran et al., 1998), enabling a self-reported measure of teachers' perceptions of their effectiveness, with a greater score indicating a higher level of perceived effectiveness. Cronbach's alpha for the teacher questionnaire indicated good reliability of the scale ($\alpha = 0.87$).

²This domain was not included in the analysis presented here, as it was not directly associated with teacher behaviours

2.6 Sample

2.6.1 Schools

Schools were varied within the sample and denoted according to the level of education provided (33.00% primary and 67.00% secondary schools). The achieved sample of primary and secondary schools was slightly over-represented in those schools with low socio-economic contexts (as indicated by Free School Meal entitlement³) but represented a range of geographical locations (29.17% urban, 60.26% suburban, 10.57% rural contexts). Consideration was also given to the number of pupils on the school roll to provide, as far as possible, a representative number of small, medium and large schools. All schools were state-funded.

2.6.2 Teachers

The teacher sample within each school was obtained on an opportunistic basis with those teachers who wanted to participate opting in voluntarily. This resulted in an achieved sample of practitioners who possessed a range of teaching experience, from newly qualified teachers to 'veteran' teachers (31+ years). The demographic data were analysed according to the length of service in the profession and these career phase groupings were selected based on Day et al.'s (2007) six phases reflecting variations in teachers' relative and perceived effectiveness. Against the national profile, the sample included a higher number of teachers in the 8–15 and 16–23 phases, and a lower number of teachers in the 0–3 and 31+ phases. The average length of experience was 17 years (Table 29.2).

Of the 312 teachers, 64.11% were female and 35.98% were male, compared to figures for England in 2021 of 72.51% female and 27.49% male (Gov.uk, 2021). The gender balance for primary school teachers (75.73% female, 24.27% male) over-represented male teachers (compared to 85.73% female, 14.27% male

Table 29.2 Teacher demographics

Career Phase	School phase				Total % of teachers	Total No of teachers
	Primary (N)		Secondary (N)			
	Male	Female	Male	Female		
Early (0–7)	6	25	8	19	18.59	58
Mid (8–23)	17	46	51	63	56.73	177
Late (24–31+)	2	7	28	40	24.68	77
Total	25	78	87	122	100	312

³Free School Meal (FSM) entitlement was used as a proxy for socio-economic context of the schools. There were four categories as follows: FSM 1 describes schools with 0–8% of pupils eligible for free school meals. This percentage rises to 9–20% for FSM 2 schools, 21–35% for FSM 3 schools and over 35% for FSM 4 schools.

nationally (Gov.uk, 2021)). However, the sample of secondary school teachers (58.40% female, 41.60% male) slightly under-represented female teachers compared with the national profile (64.60% female, 35.40% male (Gov.uk, 2021)).

2.7 Analysis Strategy

2.7.1 Initial Exploratory Analysis

For each teacher, the mean observed effectiveness was calculated from the mean scores of each of the six teacher-related domains using the data from the ICALT Lesson Observation instrument. Similarly, the perceived effectiveness mean score was calculated from the relevant items of the TASC instrument. The means and standard deviations for both observed and perceived effectiveness were compared. The means were then calculated for groups of teachers according to school phase (primary & secondary) and gender.

Both scores ranged from 1–4, with a higher number indicating a greater effectiveness score. Two null hypotheses were developed to test if there were significant differences between perceived and observed effectiveness scores. The first null hypothesis related to the entire group of teacher participants, whilst the second examined effectiveness through the lens of career phase. These were both tested using an independent samples t-test for significance. Differences in observed effectiveness and perceived effectiveness were further analysed using one-way ANOVA to test for variances within perceived and observed effectiveness.

2.7.2 Scatter Graph Analysis

Scattergrams were plotted to explore differences in teacher observed and perceived effectiveness across all six career phases. Analysis was carried out by eye to determine clusters of scores for teachers using arbitrary measures of high, intermediate and low effectiveness. Outliers were discarded from the analysis and the mean scores for both observed and perceived effectiveness then calculated for each cluster.

2.7.3 Radar Plot Analysis

The initial exploratory analysis led to a deeper investigation of observed and perceived effectiveness through the ICALT and TASC domains, using radar plots to depict the multivariate data as described by Saary (2008). The aim was to identify if variations existed in each of the six ICALT domains (excluding engagement) across the career phases of the participants. Mean averages were calculated for participants in each domain across all six career phases and presented on radar plots. Each plot examined a different career phase and consisted of six axes, depicting each of the

ICALT domains. This allowed for subtle differences in overall scores and scores for each domain to be highlighted. Points closer to the origin of the plot denote a lower observed effectiveness, whilst those further away depict greater levels of observed effectiveness. The uniformity of the hexagon shape produced describes the relative scores for each domain. For example, a profile where domain scores were of an equal magnitude would result in a perfect hexagon. When scores varied in magnitude, the hexagon shows distortions at the vertices.

The following section reports the results of these analyses, illustrating how the three research questions were addressed.

3 Results

3.1 *Is There a Difference Between Teachers' Perceived and Observed Effectiveness?*

A null hypothesis, stating that there was no significant change, on average, between a teacher's observed and perceived effectiveness was tested to explore variations in effectiveness. An independent samples t-test was conducted to look for a significant difference between observed and perceived effectiveness (Table 29.3).

The t-test showed that there is a 95% confidence level ($t(311) = 29.4, p = <0.5$) that observed effectiveness is significantly greater than perceived effectiveness in the sample of participants. This shows that teachers perceive their effectiveness to be significantly lower than it is observed to be.

3.2 *How Do Perceived and Observed Effectiveness Vary According to Teachers' Career Phase?*

To further explore variations in observed and perceived effectiveness, a second null hypothesis was tested – that there was no significant change between a teacher's observed and perceived effectiveness across the six career phases (Table 29.4).

T-tests were conducted across each career phase to test the null hypothesis. The t-tests showed that there was a 95% confidence level that observed effectiveness is significantly greater than perceived effectiveness in each of the separate career

Table 29.3 Independent t-test comparing observed effectiveness with perceived effectiveness for all participants

	Mean observed effectiveness score	Mean perceived effectiveness score	<i>N</i>	<i>t-value</i>
All participants	3.21	2.40	312	29.4

Table 29.4 Independent t-test comparing observed effectiveness with perceived effectiveness across the six career phases

Career phase	Mean average score		t-test	
	Observed effectiveness	Perceived effectiveness	N	t-value
0–3	2.33	2.36	8	0.4*
4–7	2.47	2.31	50	10.6
8–15	3.57	2.47	101	85.1
16–23	3.57	2.47	76	74.3
24–30	3.23	2.31	39	47.9
31+	2.69	2.31	38	22.9

*Indicates where the t-value was below the critical value, resulting in the null hypothesis being accepted

A chart to show perceived and observed teacher effectiveness according to career phase

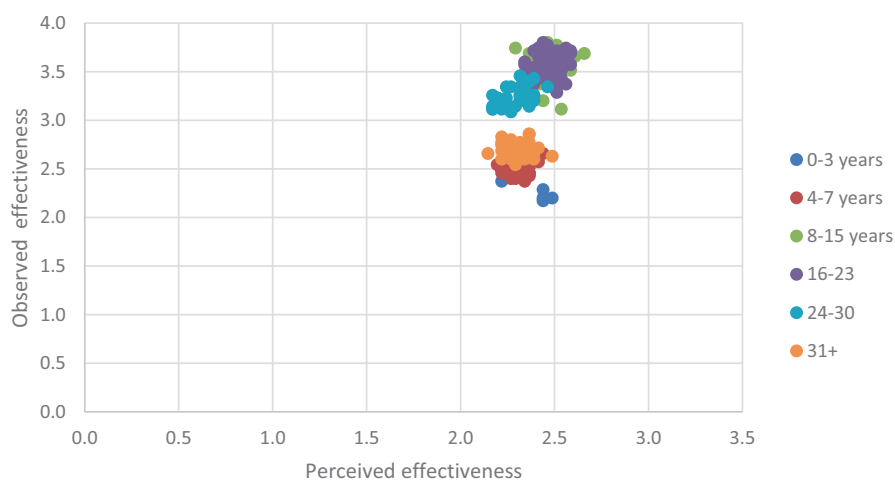


Fig. 29.1 Scattergram showing perceived and observed effectiveness across the six career phases

phases, apart from in the earliest phase (0–3 years), where there was no statistical difference.

To examine this more closely, perceived effectiveness, as measured by the questionnaire, was plotted against observed effectiveness, determined by observation. Figure 29.1 shows three clusters of data, characterised as follows:

- High perceived ($M = 2.9$) and high observed effectiveness ($M = 3.5$)
- Low perceived ($M = 2.7$) and intermediate observed effectiveness ($M = 3.2$)
- Low perceived ($M = 2.5$) and low observed effectiveness ($M = 2.5$)

The analysis identified that the early career teachers in the 0–3 phase (light blue) and 4–7 phase (orange) were situated in the low perceived and low observed effectiveness cluster. Mid-career teachers (8–15 years & 16–23 years, shown in grey and yellow respectively) were present in the high perceived and high observed effectiveness cluster. Finally, late-career teachers in the 24–30 (dark blue) and 31+ phases (green) were present in both low and intermediate observed effectiveness clusters and low perceived effectiveness. Since the clusters are distinct and represented by the majority of teachers in each of the phases, this strongly suggests that career phase may contribute to teacher effectiveness.

The variation in observed effectiveness and relatively stable scores in perceived effectiveness were further analysed using one-way ANOVA to test for variances within perceived and observed effectiveness between the six career phases (Tables 29.5 and 29.6).

The F value for perceived effectiveness scores (170.3) was above the critical value of 3.02 indicating that there were significant differences in mean perceived effectiveness across the career stages. The F value for observed effectiveness scores (1122.5), was also above the critical value, showing significant differences in mean observed effectiveness. However, the F value for observed effectiveness was far greater than the value for perceived effectiveness, indicating that whilst there was variation within perceived effectiveness scores, the variation in observed effectiveness scores was much larger.

Table 29.5 Results of one-way ANOVA test for perceived effectiveness scores across the six career phases

One-way ANOVA – Perceived effectiveness				
	Sum of squares	df	Mean square	F
Between groups	2.048	5	.410	107.261
Within groups	1.169	306	.004	
Total	3.217	311		

Table 29.6 Results of one-way ANOVA test for observed effectiveness scores across the six career phases

One way ANOVA – Observed effectiveness				
	Sum of Squares	df	Mean Square	F
Between groups	66.862	5	13.372	1122.544
Within groups	3.645	306	.012	
Total	70.507	311		

3.3 *How Can Variations in Observed Effectiveness across Teachers' Career Phases Be Explained?*

As described earlier, mean averages for each of the six ICALT domains were calculated and displayed on radar plots for each of the six career phases. Figures 29.2 a and b show these plots for the earliest career phases, 0–3 years and 4–7 years experience.

Figures 29.2 a and b illustrate the differences in observable teaching behaviours by early career teachers (phases 0–3 years and 4–7 years). It can be seen that overall, they display significantly lower overall scores for all teaching behaviours (0–3 years: $M = 2.31$, $N = 8$; 4–7 years: $M = 2.46$, $N = 50$) in comparison to both mid-career phases, (8–15 years: $M = 3.57$, $N = 101$; 16–23 years $M = 3.57$, $N = 76$) and late-career teachers, (24–30 years: $M = 3.21$, $N = 39$; 31+ years: $M = 2.68$, $N = 38$) (see Figs. 29.3 and 29.4 below for more details). The distorted hexagonal plot represents variations within the observable teaching behaviours for the early career teachers; for example, the ability of the teachers to foster a safe and stimulating learning climate and enact efficient organisation were depressed in 0–3 years in comparison with other observable indicators of teacher behaviour (see plot 2i). In the case of the 4–7 years career phase, the overall pattern was more evenly distributed as represented by a near regular hexagon, with only slight depressions visible for the same indicators.

The profile represented by the radar plots for the mid-career phases teachers (8–15 years and 16–23 years) shows considerable differences to those for the other career phases.

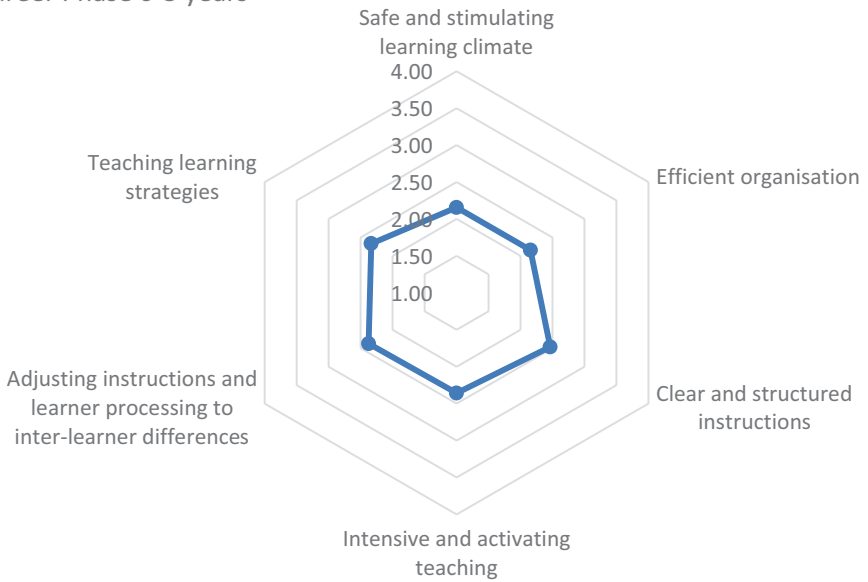
Figures 29.3 a and b illustrate the differences in observable teaching behaviours by mid-career teachers (phases 8–15 years and 16–23 years). Overall, it can be seen that they display much higher scores than those in both the earlier and later career stages (see Figs. 29.2 and 29.4) with an overall mean of 3.57 ($N = 177$). The regular hexagonal plot represents the absence of discernible variations within the highly scoring observable teaching behaviours for the middle career phases.

Again, the radar plots show teachers tend to experience another change as they enter the later career phases (24–30 years and 31+ years).

Figures 29.4 a and b illustrate the differences in observable teaching behaviours by later career teachers (phases 24–30 years and 31+ years). It can be seen that overall, they display an intermediate level of overall scoring (higher than that for the earlier career phases but lower than that for the middle career phases) for all teaching behaviours with an overall mean of 3.21 ($N = 77$).

The distorted hexagonal plots represent variations within the observable teaching behaviours for the later career teachers. For example, the ability of the teachers to provide intensive and activating teaching and adjust instructions to learners are comparatively higher within the scores of the 24–30 years (see plot 4i). Although the scores for the 31+ years teachers show the same overall profile as the 24–30 years

Career Phase 0-3 years



Career Phase 4-7 years

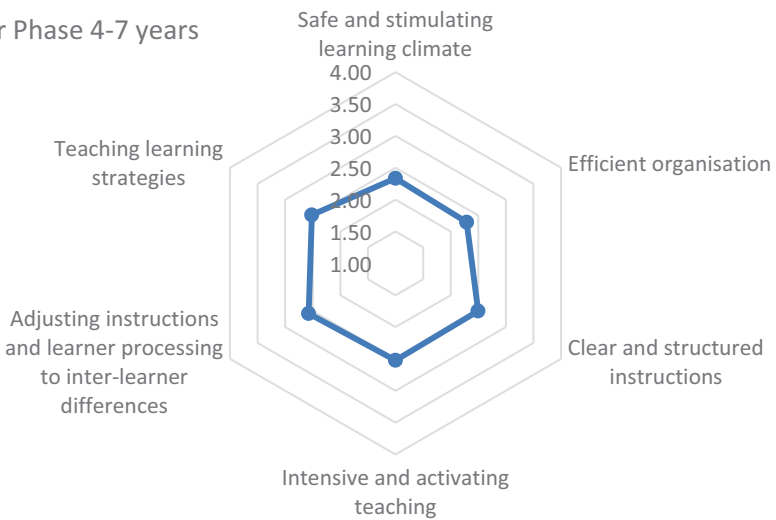
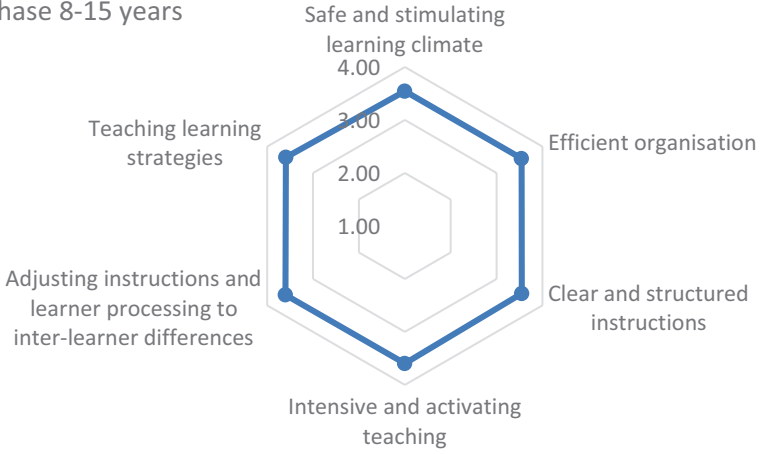


Fig. 29.2 (a) and (b) Radar plots to illustrate the differences in observed teaching behaviours according to the early career phases

Career Phase 8-15 years



Career Phase 16-23 years

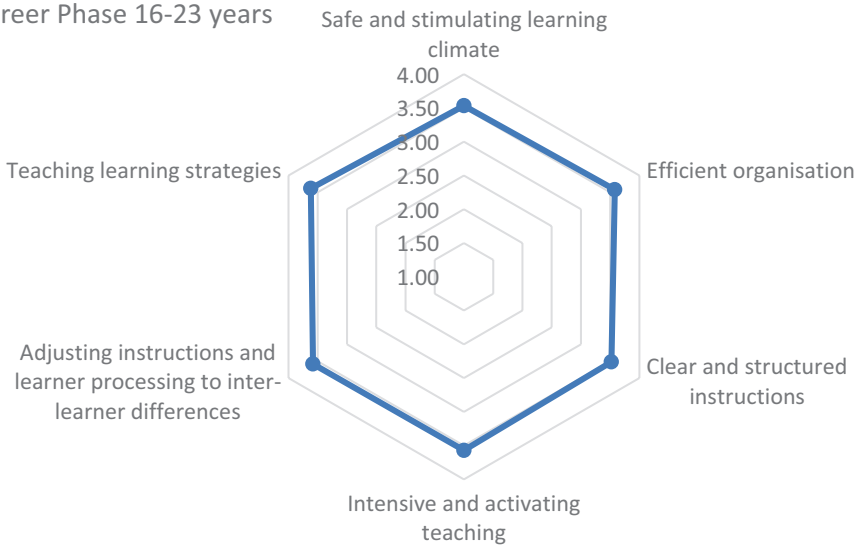
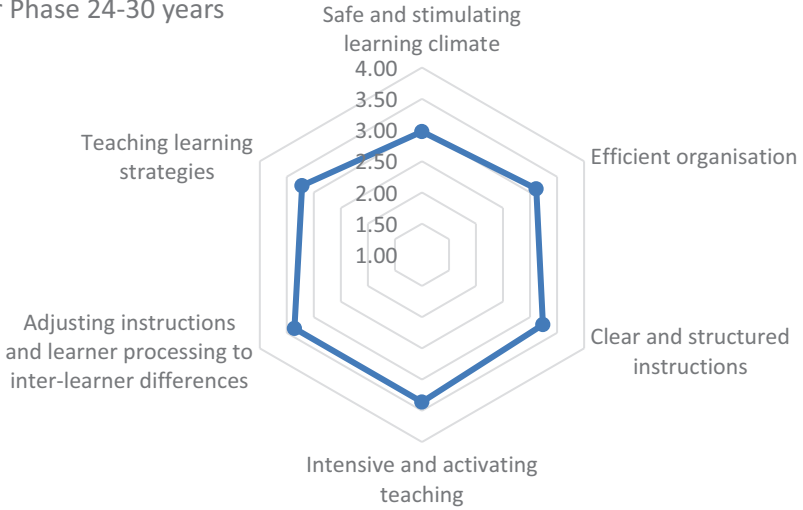
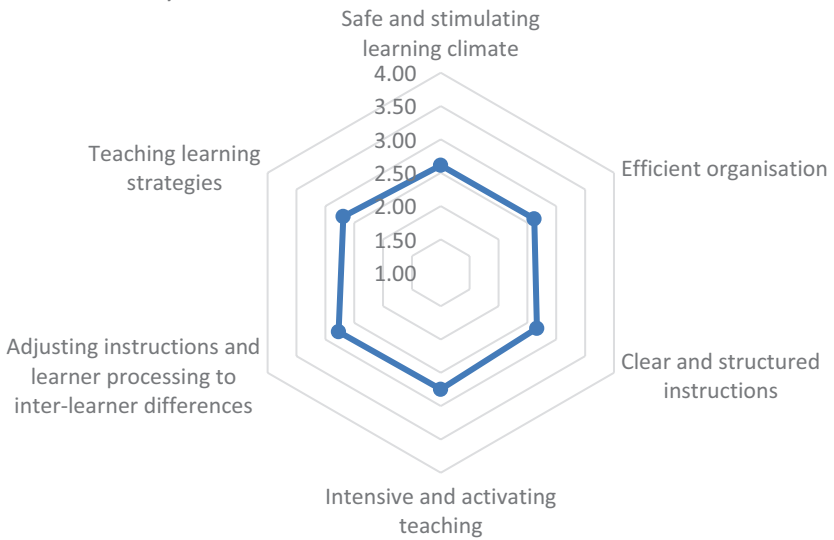


Fig. 29.3 (a) and (b) Radial plots to illustrate the differences in observed teaching behaviours according to the middle career phases

Career Phase 24-30 years



Career Phase 31+ years



Figs. 29.4 (a) and (b) Radial plots to illustrate the differences in observed teaching behaviours according to the later career phases

(as demonstrated by the same shape of an irregular hexagon), the overall scores are seen to be slightly lower (24–30 years: $M = 3.21$, $N = 39$; 31+ years: $M = 2.68$, $N = 38$).

4 Conclusions and Implications for Practice

The analysis of data reported here has shown how teachers' perceptions of their effectiveness and their observed effectiveness vary depending on the length of service. It has been shown that teachers perceive their effectiveness to be significantly lower than it is observed to be, suggesting that teachers underestimate their proficiency. This is a phenomenon that has not been explored in any depth in previous literature. In a study examining the differences between perceived and measured teaching effectiveness, Sadeghi et al. (2020) found inconsistencies between the two measures. Their study, using different instruments to measure perspectives on effectiveness for eight participants, found varied results. Two participants rated their perceived effectiveness lower than their observed effectiveness score. The remaining six participants rated themselves to be more effective than observed, highlighting the inconsistencies in self-rated measures. Conversely, results from the current study suggest that teachers consistently under-rate their performance.

When examined across the six career phases, the statistically significant differences between observed and perceived effectiveness for all participants of this study were replicated for all but one career phase. Despite all participants showing effective teaching to varying degrees, teachers in the 4–7, 8–15, 16–23, 24–30 and 31+ phases were found to perceive their effectiveness significantly lower than it was observed as being. Teachers in the earliest career phase (0–3) were found to have no statistical difference between their observed and perceived effectiveness. This could be because this group of teachers have recently completed their teacher training and are therefore more familiar with observational feedback on their effectiveness (Koni & Krull, 2018; Uhrmacher et al., 2013). However, caution is needed when interpreting the data for this group as there were only eight participants in the 0–3 career phase which might explain this anomaly.

Although the one-way ANOVA test showed statistical differences across the six career phases for both perceived and observed effectiveness, the level of significance in observed effectiveness was higher. The level of observed effectiveness rose to its highest in the mid-career phases (8–15 & 16–23), before decreasing in the late-career phases (24–30 & 31+). This could be explained by the 'disengagement stage' later career teachers have been found to experience (Day et al., 2007; Huberman, 1993; Veldman et al., 2017). T-tests confirmed this by identifying the greatest differences in *observed* versus *perceived* effectiveness for the mid-career phases (8–15 years, $t(100) = 85.1$, $p < 0.5$; 16–23 years, $t(75) = 74.3$, $p < 0.5$). It suggests that mid-career teachers' perceptions of effectiveness do not change

dramatically, unlike their observed effectiveness, which increases to the highest levels of all the career phases. Similarly, although the late-career teachers experience decreased observed effectiveness, their self-reported perceptions of effectiveness are still significantly lower (24–30 years, $t(38) = 47.9$, $p = <0.5$; 31+ years, $t(37) = 22.9$, $p = <0.5$). Interestingly, although those in the later career stages perceive their effectiveness to decline to similar levels to that of career phase 4–7 years ($M = 2.31$), their observed effectiveness remains significantly above their early career counterparts.

Examination of the six ICALT domains of observed teaching identified that whilst effective teaching behaviour increases as teachers enter the middle phases of their careers, some domains increase more rapidly than others. In the earlier career phases, there is a great variation in the scores for each domain, which is no longer present in the middle career phases. The comparatively low *efficient organisation* scores for 0–3 and 4–7 career phases indicate that these teachers are less effective at organising and structuring their lessons. Teachers in the 8–15 career phase were observed as having high levels for all six domains and this level is maintained by those in the 16–23 phase. However, this changed for teachers in the penultimate career phase (24–30), where *safe and stimulating learning climate* and *efficient organisation* scores decreased at a greater rate than the other domains. This decrease continues into the final phase (31+) when a decline in the remaining four domains is also observed.

The data suggest that *efficient organisation* is a limiting component for observed teacher effectiveness. At the start of teachers' careers, *efficient organisation* is limiting overall effectiveness. By the time a teacher is well established, *efficient organisation* rises to equal levels of *all six ICALT domains*. The decline seen in teacher effectiveness towards to end of their careers (Day et al., 2007; Huberman, 1993; Veldman et al., 2017) is shown here to be due to a decline in *safe and stimulating learning climate* and *efficient organisation*. These domains fall before the others, suggesting this drop might be a causative factor in the overall decline in observed teacher effectiveness.

In summary, the study found that there was a difference in teachers' perceived and observed effectiveness and that these appear to according to career phase. Analyses also demonstrated that these variations are associated with how teachers structure their classrooms and plan for the learning experiences of pupils. These results have implications for teachers at all stages of their career. For example, early-career teachers need to reflect on their opportunities to create and articulate explicit elements of structure and organisation within their lessons to build on increasing perceptions of classroom effectiveness. This is also crucial in retaining these practitioners in the profession. Mid-career teachers should critically engage with ways in which they can support colleagues (Lai & Lam, 2011; Mutton et al., 2011) to maintain and develop structural elements of practice, thereby affording students additional choice within lessons (Reeve & Cheon, 2021). Finally, teachers in the late-career phases could consider how to maintain a *safe and stimulating learning climate* and *efficient organisation*. Given the overall downward trajectory of effectiveness for teachers at this point in their career, professional development

could play an important factor for this group to maintain the commitment of these experienced practitioners (Brunetti & Marston, 2018).

Acknowledgements We would like to thank the schools, teachers and pupils who participated in this study. Thanks also to the University of Worcester for its continued support of the research.

References

- Aloe, A., Amo, L., & Shanahan, M. (2014). Classroom management self-efficacy and burnout: A multivariate meta-analysis. *Educational Psychology Review*, *26*, 101–126.
- Bakeman, R., & Gottman, J. (1997). *Observing interaction: An introduction to sequential analysis* (2nd ed.). Cambridge University Press.
- Baker, E., Barton, P., Darling-Hammond, L., Haertel, E., Ladd, H., & Linn, R. (2010). *Problems with the use of student test scores to evaluate teachers* [Press release]. <https://www.epi.org/publication/bp278/>
- Biesta, G. (2009). Good education in an age of measurement: On the need to reconnect with the question of purpose in education. *Educational Assessment, Evaluation and Accountability*, *21*(1), 33–46.
- Brown, S., & McIntyre, D. (1993). *Making sense of teaching*. Open University Press.
- Brown, P., Roediger, H., III, & McDaniel, M. (2014). *Make it stick: The science of successful learning*. Harvard University Press.
- Brunetti, G. J., & Marston, S. H. (2018). A trajectory of teacher development in early and mid-career. *Teachers and Teaching*, *24*(8), 874–892.
- Caprara, G., Barbaranelli, C., Borgogni, L., & Steca, P. (2003). Efficacy beliefs as determinants of Teachers' job satisfaction. *Journal of Educational Psychology*, *95*(4), 821–832.
- Caprara, G., Barbaranelli, C., Steca, P., & Malone, P. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, *44*(6), 473–490.
- Coe, R., Aloisi, C., Higgins, S., & Major, L. (2014). What makes great teaching? Review of the underpinning research. Sutton Trust. <https://dro.dur.ac.uk/13747/1/13747.pdf>
- Creemers, B., & Kyriakides, L. (2013). Using the dynamic model of educational effectiveness to identify stages of effective teaching: An introduction to the special issue. *The Journal of Classroom Interaction*, *48*(2), 4–10.
- Day, C., Stobart, G., Sammons, P., & Kington, A. (2007). Variations in the work and lives of teachers: Relative and relational effectiveness. *Teachers and Teaching: Theory and Practice*, *12*(1), 169–192.
- Gov.uk. (2021). *School workforce in England*. <https://explore-education-statistics.service.gov.uk/find-statistics/school-workforce-in-england>
- Henson, R. (2010). From adolescent angst to adulthood: Substantive implications and measurement dilemmas in the development of teacher efficacy research. *Educational Psychologist*, *37*(3), 137–150.
- Huberman, M. (1989). The professional life cycle of teachers. *Teachers College Record*, *91*(1), 31–57.
- Huberman, M. (1993). Steps toward a developmental model of the teaching career. In L. Kremer-Hayon, H. Vonk, & R. Fessler (Eds.), *Teacher professional development: A multiple perspective approach* (pp. 93–118). Swets and Zeitlinger.
- Kington, A., Sammons, P., Regan, E., Brown, E., Ko, J., & Buckler, S. (2014). *Effective classroom practice*. Open University Press.

- Klassen, R., Bong, M., Usher, E., Chong, W., Huan, V., Wong, I., & Georgiou, T. (2009). Exploring the validity of a teachers' self-efficacy scale in five countries. *Contemporary Educational Psychology, 34*(1), 67–76.
- Koni, I., & Krull, E. (2018). Differences in novice and experienced teachers' perceptions of planning activities in terms of primary instructional tasks. *Teacher Development, 22*(4), 464–480.
- Kyriacou, C. (2018). *Essential teaching skills* (4th ed.). Oxford University Press.
- Kyriakides, L., & Creemers, B. (2012). School policy on teaching and school learning environment: Direct and indirect effects upon student outcome measures. *Educational Research and Evaluation, 18*(5), 403–424.
- Labone, E. (2004). Teacher efficacy: Maturing the construct through research in alternative paradigms. *Teaching and Teacher Education, 20*(4), 341–359.
- Lai, E., & Lam, C. C. (2011). Learning to teach in a context of education reform: Liberal studies student teachers' decision-making in lesson planning. *Journal of Education for Teaching, 37*(2), 219–236.
- Little, J., Bartlett, L., Mayer, D., & Ogawa, R. (2010). The teacher workforce and problems of educational equity. *Review of Research in Education, 34*(1), 285–328.
- Maulana, R., Van de Grift, W., & Helms-Lorenz, M. (2014). Development and evaluation of a questionnaire measuring pre-service teachers' teaching behaviour: A Rasch modelling approach. *School Effectiveness and School Improvement, 26*(2), 169–194.
- Muijs, D., & Reynolds, D. (2011). *Effective teaching. Evidence and practice* (3rd ed.). SAGE.
- Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art teacher effectiveness and professional learning. *School Effectiveness and School Improvement, 25*(2), 231–256.
- Mutton, T., Hagger, H., & Burn, K. (2011). Learning to plan, planning to learn; the developing expertise of beginning teachers. *Teachers and Teaching, 17*(4), 399–416.
- OECD. (2019). *TALIS 2018 results (volume I): Teachers and school leaders as lifelong learners, TALIS*. OECD Publishing.
- OECD. (2021). *21st-century readers: Developing literacy skills in a digital world, PISA*. OECD Publishing.
- Ofsted (Office for Standards in Education, Children's Services and Skills). (2021). *Education inspection framework*. <https://www.gov.uk/government/publications/education-inspection-framework/education-inspection-framework>
- Page, L. (2016). The impact on lesson observations on practice, professionalism and teacher identity. In M. O'Leary (Ed.), *Reclaiming lesson observation: Supporting excellence in teacher learning* (pp. 62–74). Routledge.
- Reeve, J., & Cheon, S. H. (2021). Autonomy-supportive teaching: Its malleability, benefits, and potential to improve educational practice. *Educational Psychologist, 56*(1), 54–77.
- Ruddick, J., Chaplain, R., & Wallace, G. (1996). *School improvement: What can pupils tell us?* D Fulton Publishers.
- Saary, J. (2008). Radar plots: A useful way for presenting multivariate health care data. *Journal of Clinical Epidemiology, 61*(4), 311–317.
- Sadeghi, K., Richards, J. C., & Ghaderi, F. (2020). Perceived versus measured teaching effectiveness: Does teacher proficiency matter? *RELC Journal, 51*(2), 280–293.
- Sammons, P. (1999). School effectiveness: Coming of age in the 21st century. *Management in Education, 13*(5), 10–13.
- Super, D. (1957). *The psychology of careers*. Harper.
- TLRP (Teaching and learning research programme). (2013). *TLRP: Ten principles*. www.tlrp.org/themes/themes/tenprinciples.html.
- Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*(2), 202–248.
- Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783–805.

- Uhrmacher, P. B., Conrad, B. M., & Moroye, C. M. (2013). Finding the balance between process and product through perceptual lesson planning. *Teachers College Record*, *115*(7), 1–27.
- Upton, P., & Taylor, C. (2014). *Educational psychology*. Pearson.
- van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research*, *49*(2), 127–152.
- van de Grift, W. (2014). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement*, *25*(3), 295–311.
- van de Lans, R., van de Grift, W., van Veen, W., & Fokkens-Bruinsma, M. (2016). Once is not good-enough: Establishing reliability criteria for feedback and evaluation decisions based on classroom observations. *Studies in Educational Evaluation*, *50*(1), 88–95.
- Veldman, I., Admiraal, W., Mainhard, T., Wubbels, T., & van Tartwijk, J. (2017). Measuring teachers' interpersonal self-efficacy: Relationship with realized interpersonal aspirations, classroom management efficacy and age. *Social Psychology of Education*, *20*(2), 411–426.
- Wellborn, J., Connell, J., Skinner, E., & Pierson, L. (1992). *Teacher as social context (TASC). Two measures of teacher provision of involvement, structure and autonomy support*. [Technical report]. University of Rochester.
- Wragg, E. (1999). *An introduction to classroom research*. Routledge.
- Wray, D., Medwell, J., Fox, R., & Poulson, L. (2000). The teaching practices of effective teachers of literacy. *Educational Review*, *53*(1), 75–84.
- Wyatt, M. (2014). Towards a reconceptualization of teachers' self-efficacy beliefs: Tackling enduring problems with the quantitative research and moving on. *International Journal of Research and Method in Education*, *37*(2), 166–189.

Dr Benjamin Looker has worked in secondary education, holding a variety of positions, including assistant headteacher for teaching and learning. Now working as a Principal Lecturer, Ben has developed his passion for research in education. His educational research interests are focused on social psychology of education, including examining the various manifestations of alienation pupils might experience while at secondary school. Having a background in natural sciences, but now researching in social sciences, Ben is particularly interested in the intersection of these two research disciplines and has formulated a critical realist approach to grounded theory.

Prof. Alison Kington is a Professor in Psychology of Education at the University of Worcester, UK, and a Chartered Psychologist. She has worked in a number of research and teaching roles and gained extensive experience of, and expertise in, designing and conducting mixed methods research in education and social psychology. Her research focuses on the nature, quality and dynamics of educational relationships and she has a particular interest in the influence of teacher identity and career phase on classroom relationships (adult-child & peer) and interactions. Alison has led a range of international and national research projects funded by Research Councils and Government agencies and has published widely in relation to her substantive and methodological interests.

Kimberley Hibbert-Mayne has worked in education since 2006. She worked as a Physical Education (PE) teacher in secondary schools before embarking on a career in teacher education, taking on the role of PGCE Secondary tutor and lead for the PE and Professional Studies programmes at the University of Worcester, UK. Her MA in Education and other research projects have predominantly focused on areas of social psychology. Kim is particularly interested in how an individual's attitudes, values and personality traits affect their experiences and behaviours within the teaching profession. She is passionate about preparing trainee teachers for a long, happy and healthy career.

Dr. Karen Blackmore Originally a bio-pharmaceutical scientist, Karen has taught in a range of schools and university academic departments for over 20 years. Her present position is Senior Lecturer in Science Education at the University of Worcester, where she uses her expertise to lead the primary science initial teacher education provision. Karen has led a number of classroom-based empirical research projects, including the design and evaluation of innovative technology-enhanced pedagogy. Her publications in this area have been used as a basis for research-informed teaching with her student teachers. She is also interested in teacher professional identity and efficacy and the impact of these on the development of effective social and learning relationships.

Dr. Scott Buckler has an extensive career in education, as a primary and secondary school teacher, e-learning developer, and as a Principal Lecturer, having worked for four universities predominantly in education, psychology and inclusion. In recent years, Scott has returned to school teaching to refresh his practical classroom experience and has been awarded Chartered Teacher Status. Scott has a PhD in anthropology and is widely published in the areas of psychology and education. He is a Chartered Psychologist with expertise in transpersonal psychology and applied educational psychology.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Part V

Effective Teaching in Complex Environments: Differentiation and Adaptive Teaching

Part V Overview

This part presents studies that shed light on differentiated instruction from different perspectives and stakeholders in education. The authors contributing to this part of the book are associated with universities in Flanders (Belgium), the Netherlands and South Africa.

A number of chapters report studies regarding questions how differentiated instruction can be improved. The promising approach of Lesson Study (described in Chap. 32) that focuses explicitly on students' learning, is expected to improve the supportive capacity of teachers to meet student needs more effectively. They argue that this assumption should be tested in the future. Another promising approach is introduced in Chap. 34 regarding teachers' reflections on classroom interactions, revealing a pattern of high teacher activity and low student activity, to improve a focus on student learning to promote effective teaching. The study presented in Chap. 33 reveals the essence of eliciting evidence of learning during the lesson, as an extra phase of differentiated instruction that encompasses the (pro-active) lesson preparation, (inter-action) execution and the (retro-active) reflection on the lesson. Chapter 30 highlights the importance of teachers' philosophical stance to implementing differentiated instruction, the importance of perceiving and implementing differentiated instruction as a pedagogical model, and the importance and complexity of professional development with regard to differentiated instruction.

In a study conducted in the Gauteng province of South Africa (Chap. 31) it was found that teachers were not always aware of students' needs in the classroom and the challenges that impede their effective learning. The possible reasons could be inadequate training of teachers to identify students' learning barriers and to create and implement differentiated activities; teachers experiencing a lack of time to complete the curriculum; a lack of resources; teaching large classes; and an inability to manage and maintain discipline in classes.

The study reported in Chap. 35 focusses on DI in primary education, revealing that teachers generally monitor student achievement. Although efforts are made to

adapt instructions, high-achievers are rarely considered in these practices. The flexibility of within-class grouping and refining student-need diagnostic strategies deserve more attention. Chapter 36 reports on teachers' intentions (towards students with and without special education needs (SEN)) to differentiate instruction in regular secondary vocational education. Additionally, one-to-one classroom interactions between teachers and students with and without special educational needs were analyzed.

Chapter 30

Differentiated Instruction as an Approach to Establish Effective Teaching in Inclusive Classrooms



Esther Gheysens , Júlia Griful-Freixenet , and Katrien Struyven 

Abstract Differentiated Instruction has been promoted as a model to create more inclusive classrooms by addressing individual learning needs and maximizing learning opportunities. Whilst differentiated instruction was originally interpreted as a set of teaching practices, theories now consider differentiated instruction rather a pedagogical model with philosophical and practical components than the simple act of differentiating. However, do teachers also consider differentiated instruction as a model of teaching? This chapter is based on a doctoral thesis that adopted differentiated instruction as an approach to establish effective teaching in inclusive classrooms. The first objective of the dissertation focused on how differentiated instruction is perceived by teachers and resulted in the DI-Quest model. This model, based on a validated questionnaire towards differentiated instruction, pinpoints different factors that explain differences in the adoption of differentiated instruction. The second objective focused on how differentiated instruction is implemented. This research consisted of four empirical studies using two samples of teachers and mixed method. The results of four empirical studies of this dissertation are discussed and put next to other studies and literature about differentiation. The conclusions highlight the importance of teachers' philosophy when it comes to implementing differentiated instruction, the importance of perceiving and implementing differentiated instruction as a pedagogical model and the importance and complexity of professional development with regard to differentiated instruction.

E. Gheysens (✉)
Ghent University, Ghent, Belgium

Vrije Universiteit Brussel, Aalst, Belgium
e-mail: esther.gheysens@ugent.be

J. Griful-Freixenet
Vrije Universiteit Brussel, Aalst, Belgium
Universitat de Barcelona, Barcelona, Spain

K. Struyven
Vrije Universiteit Brussel, Aalst, Belgium
Universiteit Hasselt, Hasselt, Belgium

Keyword Differentiated instruction · Effective teaching · Inclusive classrooms

1 Introduction

Differentiated Instruction (DI) has been promoted as a model to facilitate more inclusive classrooms by addressing individual learning needs and maximizing learning opportunities (Gheysens et al., 2020c). DI aims to establish maximal learning opportunities by differentiating the instruction in terms of content, process, and product in accordance with students their readiness, interests and learning profiles (Tomlinson, 2017). This chapter is based on a doctoral thesis that adopted DI as an approach to establish effective teaching in inclusive classrooms. This doctoral dissertation consisted of four empirical studies towards the conceptualisation and implementation of DI (Gheysens, 2020). This chapter summarizes the most important results of this dissertation and includes three parts. First the conceptualisation of DI is discussed. Second, we discuss literature findings regarding the effectiveness of DI. Third, the results of the studies about the implementation of DI are discussed. Finally, based on the previous parts some recommendations for implementation are presented.

2 Conceptualisation of Differentiated Instruction

2.1 *Defining Differentiated Instruction*

Differentiated instruction (DI) is an approach that aims to meet the learning needs of all students in mixed ability classrooms by establishing maximal learning and differentiating instruction with regard to content, process and product in accordance with student needs in terms of their readiness (i.e., student's proximity to specified learning goal), interests (i.e., passions, affinities that motivate learning) and learning profiles (i.e., preferred approaches to learning) (Tomlinson, 2014). Whilst DI was originally interpreted as a set of teaching practices or simplified as the act of differentiating (e.g. van Kraayenoord, 2007; Tobin, 2006), it is evolved towards a pedagogical model with philosophical and practical components (Gheysens, 2020). This model is rooted in the belief that diversity is present in every classroom and that teachers should adjust their education accordingly (Tomlinson, 1999). Tomlinson (2017) states that DI is an approach where teachers are proactive and focus on common goals for each student by providing them with multiple options in anticipation of and in response to differences in readiness, interest, and learning needs (Tomlinson, 2017). From this perspective, differentiation refers to an

educational process where students are made accountable for their abilities, talents, learning pace, and personal interests (Op 't Eynde, 2004). This means that teachers proactively plan varied activities addressing what students need to learn, how they will learn it, and how they show what they have learned. This increases the likelihood that each student will learn as much as he or she can as efficiently as possible (Tomlinson, 2005). Moreover, DI emphasizes the needs of both advanced and struggling learners in mixed-ability classroom. In more detail, Bearne (2006) and Tomlinson (1999) consider differentiation as an approach to teaching in which teachers proactively adjust curricula, teaching methods, resources, learning activities and student product so that various student's needs are satisfied (individuals or small groups) and every student is provided with maximum learning opportunities (in Tomlinson et al., 2003).

2.2 *The DI-Quest Model*

Considering DI as a pedagogical model rather than as a set of teaching strategies became also clear in the validity study of Coubergs et al. (2017) when they tried to measure DI empirically. Their research resulted in the so-called 'DI-Quest model', based on the DI-questionnaire the researchers developed for investigating DI. This model pinpoints different factors that explain differences in the adoption of differentiated instruction (Coubergs et al., 2017). It was inspired by the differentiated instruction model developed by Tomlinson (2014), which presents a step by step process demonstrating how a teacher moves from thinking about DI toward implementing it in the classroom. According to this model, the teacher can differentiate content, process, product, and environments to respond to different needs in learning based on students' readiness, learning profiles, and interests. Tomlinson (2014) also stipulates that, to respond adequately to students' learning needs, teachers should apply general classroom principles such as respectful tasks, flexible grouping, and ongoing assessment and adjustment. In contrast with Tomlinson's well-known DI model, which also contains concepts relating to good teaching, the DI-Quest model distinguishes teachers who use DI less often from those who use it more often (Gheysens et al., 2020c). The DI-Quest model comprises five factors. The five factors are presented in three categories. The key factor, similar to Tomlinson's (2014) model, is adapting teaching to students' readiness, interests, and learning profiles. This is the main factor because it represents the 'core business' of differentiating: the teachers adapt his/her teaching to three essential differences in learning. The second and third factors represent DI as a philosophy. The fourth and fifth factor represent differentiated strategies in the classroom (Fig. 30.1). Below the figure the different factors are discussed on detail.

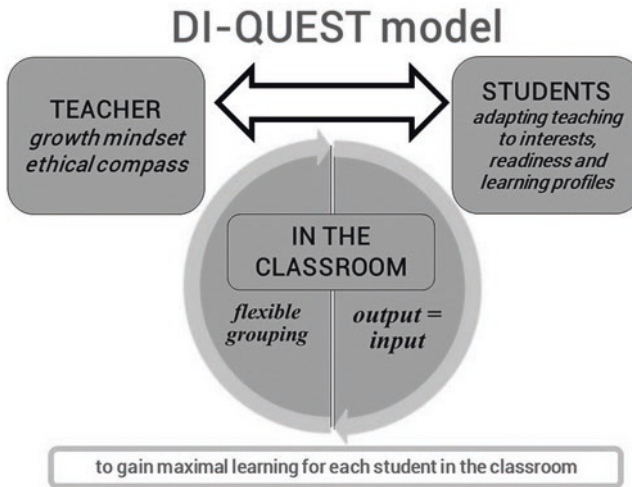


Fig. 30.1 The DI-Quest model

2.2.1 Adaptive Teaching

Adaptive teaching illustrates that the teacher provides various options to enable students to acquire information, digest, and express their understanding in accordance with their readiness, interests, and learning profiles (Tomlinson, 2001). Differences in learning profiles are described by Tomlinson and colleagues (2003, p. 129) as “a student’s preferred mode of learning that can be affected by a number of factors, including learning style, intelligence preference and culture.” Applying different learning profiles positively influences the effectiveness of learning because students get the opportunity to lean the way they learn best. Responding to student interests also appears to be related to more positive learning experiences, both in the short and long term (Woolfolk, 2010; Tomlinson et al., 2003). Ryan and Deci (2000) claimed that understanding what motivates students will help develop interest, joy, and perseverance during the learning process. Thus, investing in differences in interests increases learning motivation among students. Taking account of students’ readiness can also lead to higher academic achievement. Readiness focuses on differences arising from a student’s learning position relative to the learning goals that are to be attained (Woolfolk, 2010). When taking students’ readiness into account enables every student to attain the learning objectives in accordance with their learning pace and position (Gheysens et al., 2021).

2.2.2 Philosophy of DI

The first philosophical factor to consider is the ‘growth mindset’. Tomlinson (2001) addressed the concept of mindset in her DI model by stating that a teacher’s mindset can affect the successful implementation of differentiated instruction (Sousa &

Tomlinson, 2011). Teachers with a growth mindset set high goals for their students and believe that every student is able to achieve success when they show commitment and engagement (Dweck, 2006). The second philosophical factor is the 'ethical compass'. This envisions the use of curriculum, textbooks, and other external influences as a compass for teaching rather than observations of the student (Coubergs et al., 2017; Tomlinson & Imbeau, 2010). An ethical compass that focuses on the student embodies the development of meaningful learning outcomes, devises assessments in line with these, and creates engaging lesson plans designed to enhance students' proficiency in achieving their learning goals (Tomlinson & Imbeau, 2010). Research on self-reported practices demonstrated that teachers with an overly rigid adherence to a curriculum that does not take students' needs into account, report to adopt less adaptive teaching practices (Coubergs et al., 2017; Gheysens et al., 2020c).

2.2.3 Differentiated Classroom Practices

The next factor is the differentiated practice to be explained is 'flexible grouping'. Switching between homogeneous and heterogeneous groups helps students to progress based on their abilities (when in homogeneous groups) and facilitates learning through interaction (when in heterogeneous groups) (Whitburn, 2001). Given that the aim of differentiated instruction is to provide maximal learning opportunities for all students, variation between homogeneous and heterogeneous teaching methods is essential. Coubergs et al. (2017) found that combining different forms of flexible grouping positively predicts the self-reported use of adaptive teaching in accordance with differences in learning. The final factor in the DI-Quest model is the differentiated practice 'Output = input.' This factor represents the importance of using output from students (such as information from conversations, tasks, evaluation, and classroom behaviour) as a source of information. This output of students is input for the learning process of the students themselves by providing them with feedback. But this output is also crucial input for the teacher in terms of information about how students react to his/her teaching (Hattie, 2009). Assessment and feedback are not the final steps in the process of teaching, but they are an essential part of the process of teaching and learning (Gijbels et al., 2005). In this regard, Coubergs et al. (2017) state that including feedback as an essential aspect of learning positively predicts the self-reported use of adaptive teaching.

3 Effectiveness of Differentiated Instruction

Several studies dealing with the effectiveness of DI have demonstrated a positive impact on student achievement (e.g. Beecher & Sweeny, 2008; Endal et al., 2013; Mastropieri et al., 2006; Reis et al., 2011; Smale-Jacobse et al., 2019; Valiandes, 2015). However, while recent theories plead for a more holistic interpretation of DI,

being a philosophy and a practice of teaching, empirical studies on the impact on student learning are often limited to one aspect of DI, e.g. ability grouping, tiering, heterogenous grouping, individualized instruction, mastery learning or another specific operationalization of DI (e.g. Bade & Bult, 1981; Tomlinson, 1999; Vanderhoeven, 2004; Smale-Jacobse et al., 2019). Often studies on DI are also fragmented in studies on ability grouping, tiering, heterogenous grouping, individualized instruction, mastery learning or another specific operationalization of DI (Coubergs et al., 2013; Smale-Jacobse et al., 2019). Although effectiveness can be found for most of these operationalisations, overall the evidence is limited and sometimes even inconclusive (e.g. evidence of the benefits on ability grouping). Indeed research indicates that DI has the power to benefit students' learning. However, this might not always be the case for all students. For example Reis and colleagues demonstrated that at-risk students are most likely to benefit from DI (e.g. Reis et al., 2011). By contrast, experimental research on DI by Valiandes (2015) showed that although the socioeconomic status of students correlated with their initial performance, it had no effect on their progress. This confirmed that DI can maximize learning outcomes for all students regardless of their socioeconomic background. It also depends on how DI is implemented, for example the effects of ability grouping may differ for subgroups of students (Coubergs et al., 2013). A recent review on DI concluded that studies of effectiveness of DI overall report small to medium-sized positive effects of DI on student achievement. However, the authors of this study plead for more empirical studies towards the effectiveness of DI on both academic achievement and affective students' outcomes, such as attitudes and motivation (Smale-Jacobse et al., 2019).

4 Implementation of Differentiated Instruction

Differentiated instruction is often presented in a fragmented fashion in studies. For example, it can be defined as a specific set of strategies (Bade & Bult, 1981; Woolfolk, 2010) or studies with regard to the effectiveness of DI often focus on specific differentiated classroom actions, rather than on DI as a whole-classroom approach (Smale-Jacobse et al., 2019). Moreover, DI is not only in studies fragmented defined and investigated, DI is also perceived by teachers in a fragmented way (Gheysens, 2020). For example, using mixed methods, this study explored to what degree differentiated practices are implemented by primary school teachers in Flanders (Gheysens et al., 2020a). Data were gathered by means of three different methods, which are compared: teachers' self-reported questionnaires (N = 513), observed classroom practices and recall interviews (N = 14 teachers). The results reveal that there is not always congruence between the observed and self-reported practices. Moreover, the study seeks to understand what encourages or discourages teachers to implement DI practices. It turns out that concerns about the impact on students and school policy are referred to by teachers as impediments when it comes to adopting differentiated practices in classrooms. On teacher level, some teachers

expressed a feeling of powerlessness towards their teaching and have doubts if their efforts are good enough. On school level, a development plan was often missing which gave teachers the feeling that they are standing alone (Gheysens et al., 2020a). Other studies confirm that when beliefs about teaching and learning are different among various actors involved in a school, this can limit DI implementation (Beecher & Sweeny, 2008). However, we know from the DI-Quest model how important a teachers' mindset is when it comes to implementing DI. In this specific study teachers were asked about both hindrances and encouragements to implement DI. Teachers only responded with hindrances. In addition, flexible grouping, which in theory is an ideal teaching format when it comes to differentiation, occurs often randomly in the classroom without the intention to differentiate. The researchers of this study concluded that teachers do not succeed in implementing DI to the fullest because their mindset about DI is not as advanced as their abilities to implement differentiated practices. These practices, such as flexible grouping for example, are often part of the curriculum. Moreover, also in teacher education programmes pre-service teachers are trained to use differentiated strategies. However, teacher education programmes approach DI mostly again as a set of teaching practices. Teaching a mindset is much more difficult and complicated. This focus on DI as only a practice and as a pedagogical model, like the DI-Quest model demonstrates, leads to partial implementation of DI. DI is then perceived as something teachers can do "sometimes" in their classrooms, rather than a pedagogical model that is embedded in the daily teaching and learning process (Gheysens et al., 2020a).

In other words, one aspect of DI is often implemented, one specific teaching format is applied, or one strategy is adopted to deal with one specific difference between learning. As a consequence, some aspects will be improved or some students will benefit from this approach, but the desired positive effects on the total learning process of all the students that theories about DI promise, will remain unforthcoming. Below some recommendations are listed to implement DI more as a pedagogical model and less fragmented.

4.1 Importance of the Teachers' Philosophy

Review studies which investigated the effectiveness and implementation of specific operationalizations of DI (for example grouping) report small to medium effects on student achievement (Coubergs et al., 2013; Smale-Jacobse et al., 2019). Although theories recommend approaching DI as a holistic concept, the effectiveness of such a holistic approach on student learning has, to our knowledge, not yet been investigated. We emphasize the importance of presenting and perceiving DI as a pedagogical model that is regarded as a philosophy of teaching and a collection of teaching practices (Tomlinson, 2017). Thus, DI is considered a pedagogical model that is influenced by teachers' mindset and one which encourages teachers to be proactive, involves modifying curricula, teaching methods, resources, learning activities and student products in anticipation of, and response to, student differences in readiness,

interests and learning profiles, in order to maximize learning opportunities for every student in the classroom (Coubergs et al., 2017; Tomlinson, 2017). In this regard we would also like to emphasize that these modifications do not necessarily involve new teaching strategies and extra workload for teachers, but require that teachers shift their mindset and start acting more pro-actively, planned better and be more positive. In a study that investigated the effectiveness of a professional development programme about inclusive education on teachers' implementation of differentiated instruction, teachers stated that after participating in the programme they did not necessarily adopt more differentiated practices, but they did the ones they used more thoroughly (Gheysens et al., 2020b). As demonstrated in the DI-Quest model, in order to implement DI as a pedagogical model, it is essential to start with the teachers' philosophy. However, changing a philosophy does not come about overnight, but rather demands time and patience (Gheysens, 2020).

4.2 Importance and Complexity of Professional Development

When DI becomes a pedagogical model that consists of both philosophy and practice components, and furthermore demands that teachers have a positive mindset towards DI in order to implement DI effectively, professional development for some teachers is necessary to strengthen their competences and to support them in embedding DI in their classrooms. Depending on the current mindset of the teacher, some will need more support, while for other teachers differentiating comes naturally. However, if we want teachers to implement DI as a pedagogical model and not just as fragmented practices, teachers need to be prepared and supported. Professional development is essential for teachers to respond adequately to the changing needs of students during their careers (Keay & Lloyd, 2011; EADSNE, 2012). However, professional development is also complex. The final study in the dissertation of Gheysens (2020) investigated the effectiveness of a professional development programme (PDP) aimed at strengthening the DI competences of teachers. A quasi-experimental design consisting of a pre-test, post-test, and control group was used to study the impact of the programme on teachers' self-reported differentiated philosophies and practices. Questionnaires were collected from the experimental group ($n = 284$) and control group ($n = 80$) and pre- and post-test results were compared using a repeated measures ANOVA. Additionally, interviews with a purposive sample of teachers ($n = 8$) were conducted to explore teachers' experiences of the PDP. The results show that the PDP was not effective in changing teachers' DI competences. Multiple explanations are presented for the lack of improvement such as treatment fidelity, the limitations of instruments, and the necessary time investment (Gheysens et al., 2020b).

We found similar information in other studies. For example Brighton et al. (2005) stated that the biggest challenge for most teachers is that DI questions their previous beliefs. This ties in with our emphasis on teachers' mindset. To participate in professional development, teachers need to have/keep an open mind in order to respond to

new forms of diversity and new opportunities for collaborating with colleagues. Although continued professional development is necessary and important for teachers, it is a complex process. We refer to the work of Merchie et al. (2016) who identified nine characteristics of effective professional development, with one of them being that the supervisor is of high quality and is competent when it comes to giving and receiving constructive feedback and imparting other coaching skills (Merchie et al., 2016). Literature states that professional development is only successful if teachers are active participants, if they have a voice in what and how they learn things, and if the PDP is tailored to the specific context (Merchie et al., 2016). However, PDP often works towards a specific goal which is not always very flexible. A suitable coach is able to find a balance between these two extremes. Or, specifically within inquiry-based learning as an example, the coach needs to find the fragile balance between telling the teachers what to do, and letting them find their own answers. Finding such a balance and guiding teachers towards looking for and finding the answers they need is important if we wish to establish the desired improvement we want to see in teachers' professional development. In this regard, Willegems et al. (2016) plead for the role of a broker as a bridge-maker in professional development trajectories, in addition to the role of coach (Willegems et al., 2016).

4.3 Importance of Collaboration

In addition, collaboration is indeed essential for effective professionalisation (Merchie et al., 2016) and beneficial for DI implementation (De Neve et al., 2015; Latz & Adams, 2011). In a professional development study where inquiry-based learning was applied to teams of teachers at schools, teachers reported positive experiences in discussing their individual learning activities, and during the programme became aware of the need to work together on the collective development of knowledge in the school. They all agreed that to implement DI they needed to collaborate more. A common school vision and policy is necessary for the implementation of specific differentiated measures, as these currently differ between teachers and grades, and can be confusing for students. This is consistent with previous research that states that collaboration is crucial for creating inclusive classrooms (Hunt et al., 2002; Mortier et al., 2010; EADSNE, 2012; Claasen et al., 2009; Mitchel, 2014). A first step in this process is realising that collaboration is beneficial for both teachers and students (EADSNE, 2012).

5 Conclusion

The chapter summarizes a doctoral dissertation that started with the assumption from theory that differentiated instruction can be adopted to create more inclusive classrooms. Theories describe DI as both a teaching practice and a philosophy, but

the concept is rarely measured as such. Empirical evidence about the effectiveness and operationalisation of differentiating is limited. The general aim of this research was to gain a more in-depth understanding of the concept of DI. This main aim was subdivided into two objectives. The first objective focused on how DI is perceived by teachers and resulted in the DI-Quest model. The second objective focused on how DI is implemented. Four empirical studies were conducted to address these objectives. Two different samples spread over three years were adopted (1302 teachers in study 1 and 1522 teachers in studies 2, 3 and 4) and mixed methods were applied to investigate these research goals. In this chapter the results of these studies were put next to other studies and literature about differentiation. The conclusions highlight the importance of teachers' philosophy when it comes to implementing DI, the importance of perceiving and implementing DI as a pedagogical model and the importance and complexity of professional development with regard to DI. Overall, the authors of this dissertation conclude that DI can be as promising as theories say when it comes to creating inclusive classrooms, but at the same time their research illustrated that the reality of DI in classrooms, is far more complex than the theories suggest.

References

- Bade, J., & Bult, H. (1981). *De praktijk van interne differentiatie. Handboek voor de leraar*. Uitgeverij Intro Nijkerk.
- Bearne, E. (2006). *Differentiation and diversity in the primary school*. Routledge.
- Beecher, M., & Sweeny, S. M. (2008). Closing the achievement gap with curriculum enrichment and differentiation: One school's story. *Journal of Advanced Academica*, 19(3), 502–530.
- Claasen, W., de Bruïne, E., Schuman, H., Siemons, H. & van Velthooven, B. (2009). *Inclusief bekwaam. Generiek competentieprofiel inclusief onderwijs - LEOZ Deelproject 4*. Garant.
- Brighton, C. M., Hertberg, H. L., Moon, T. R., Tomlinson, C. A., & Callahan, C. M. (2005). *The feasibility of high-end learning in a diverse middle school*. National Research Center on the Gifted and Talented.
- Coubergs, C., Struyven, K., Engels, N., Cools, W., & De Martelaer, K. (2013). *Binnenklasdifferentiatie. Leerkansen voor alle leerlingen*. Acco.
- Coubergs, C., Struyven, K., Vanthournout, G., & Engels, N. (2017). Measuring teachers' perceptions about differentiated instruction: The DI-Quest instrument and model. *Studies in Educational Evaluation*, 53, 41–54.
- De Neve, D., Devos, G., & Tuytens, M. (2015). The importance of job resources and self-efficacy for beginning teachers' professional learning in differentiated instruction. *Teaching and Teacher Education*, 47, 30–41.
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.
- EADSNE (European Agency for Development in Special Needs Education). (2012). *Lerarenopleiding en inclusie. Profiel van inclusieve leraren*. European Agency for Development in Special Needs Education.
- Endal, G., Padmadewi, N., & Ratminingsih, M. (2013). The effect of differentiated instruction and achievement motivation on students' writing competency. *Journal Pendidikan Bahasa Inggris*, 1.
- Gheysens, E. (2020). *Adopting differentiated instruction to create inclusive classrooms*. Crazy Copy Center Productions.

- Gheysens, E., Consuegra, E., Vanslambrouck, S., Engels, N., & Struyven, K. (2020a). Differentiated instruction in practice: Do teachers walk the talk? *Pedagogische Studieën*, 97, 163–186.
- Gheysens, E., Consuegra, E., Engels, N., & Struyven, K. (2020b). Good things come to those who wait: The importance of professional development for the implementation of differentiated instruction. *Frontiers in Education*, 5, 96.
- Gheysens, E., Coubergs, C., Griful-Freixenet, J., Engels, N., & Struyven, K. (2020c). Differentiated instruction: The diversity of teachers' philosophy and practice to adapt teaching to students' interests, readiness and learning profiles. *International Journal of Inclusive Education*, 26, 1383.
- Gheysens, E., Consuegra, E., Engels, N., & Struyven, K. (2021). Creating inclusive classrooms in primary and secondary schools: From noticing to differentiated practices. *Teaching and Teacher Education*, 100, 103210.
- Gijbels, D., Dochy, F., Van den Bossche, P., & Segers, M. (2005). Effects of problem-based learning: A meta-analysis from the angle of assessment. *Review of Educational Research*, 75(1), 27–61.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Taylor & Francis Ltd.
- Hunt, P., Soto, G., Maier, J., Müller, E., & Goetz, L. (2002). Collaborative teaming to support students with augmentative and alternative communication needs in general education classrooms. *AAC: Augmentative and Alternative Communication*, 18(1), 20–35.
- Keay, J. K., & Lloyd, C. M. (2011). Developing inclusive approaches to learning and teaching. In J. K. Keay & C. M. Lloyd (Eds.), *Linking children's learning with professional learning. Impact, evidence and inclusive practice* (pp. 31–44). Sense Publishers.
- Latz, A. O., & Adams, C. M. (2011). Critical differentiation and the twice oppressed: Social class and giftedness. *Journal for the Education of the Gifted*, 34(5), 773–789.
- Mastropieri, M. A., Scruggs, T. E., Norland, J. J., Berkeley, S., McDuffie, K., Tornquist, E. H., & Connors, N. (2006). Differentiated curriculum enhancement in inclusive middle school science effects on classroom and high-stakes tests. *The Journal of Special Education*, 40(3), 130–137.
- Merchie, E., Tuytens, M., Devos, G., & Vanderlinde, R. (2016). Evaluating teachers' professional development initiatives: Towards an extended evaluative framework. *Research Papers in Education*, 33, 1–26.
- Mitchel, D. (2014). *What really works in special and inclusive education. Using evidence-based teaching strategies* (2nd ed.). Routledge.
- Mortier, K., Van Hove, G., & De Schauwer, E. (2010). Supports for children with disabilities in regular education classrooms: An account of different perspectives in Flanders. *International Journal of Inclusive Education*, 14(6), 543–561.
- Op 't Eynde. (2004). Leren doe je nooit alleen: differentiatie als een sociaal gebeuren. *Impuls*, 35(1), 4–13.
- Reis, S. M., McCoach, D. B., Little, C. A., Muller, L. M., & Kaniskan, R. B. (2011). The effects of differentiated instruction and enrichment pedagogy on reading achievement in five elementary schools. *American Educational Research Journal*, 48(2), 462–501.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and Well-being. *American Psychologist*, 55(1), 68.
- Smale-Jacobse, A. E., Meijer, A., Helms-Lorenz, M., & Maulana, R. (2019). Differentiated instruction in secondary education: A systematic review of research evidence. *Frontiers in Psychology*, 10, 2366.
- Sousa, D. A., & Tomlinson, C. A. (2011). *Differentiation and the brain: How neuroscience supports the learner – Friendly classroom*. Solution Tree Press.
- Tobin, R. (2006). Five ways to facilitate the teacher assistant's work in the classroom. *Teaching Exceptional Children Plus*, 2(6), n6.
- Tomlinson, C. A. (1999). *The differentiated classroom. Responding to the needs of all learners*. ASCD.

- Tomlinson, C. A. (2001). *How to differentiate instruction in mixed-ability classrooms*. Association for Supervision and Curriculum Development.
- Tomlinson, C. A. (2005). Grading and differentiation: Paradox or good practice? *Theory Into Practice*, 44(3), 262–269.
- Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners*. ASCD.
- Tomlinson, C. A. (2017). *How to differentiate instruction in academically diverse classrooms*. ASCD.
- Tomlinson, C. A., & Imbeau, M. B. (2010). *Leading and managing a differentiated classroom*. ASCD.
- Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., Conover, L. A., & Reynolds, T. (2003). Differentiating instruction in response to student readiness, interest and learning profile in academically diverse classrooms: A review of the literature. *Journal for the Education of the Gifted*, 27(2/3), 119–145.
- Valiandes, S. (2015). Evaluating the impact of differentiated instruction on literacy and reading in mixed ability classrooms: Quality and equity dimensions of education effectiveness. *Studies in Educational Evaluation*, 45, 17–26.
- van Kraayenoord, C. E. (2007). School and classroom practices in inclusive education in Australia. *Childhood Education*, 83(6), 390–394.
- Vanderhoeven, J. L. (2004). *Positief omgaan met verschillen in de leeromgeving. Een visie op differentiatie en gelijke kansen in authentieke middenscoles*. Uitgeverij Antwerpen.
- Whitburn, J. (2001). Effective classroom organisation in primary schools: Mathematics. *Oxford Review of Education*, 27(3), 411–428.
- Willegems, V., Consuegra, E., Struyven, K., & Engels, N. (2016). How to become a broker: The role of teacher educators in developing collaborative teacher research teams. *Educational Research and Evaluation*, 22(3–4), 173–193.
- Woolfolk, A. (2010). *Educational psychology*. The Ohio State University: Pearson Education International.

Esther Gheysens obtained her doctoral degree at the Vrije Universiteit Brussel in 2020, with a dissertation titled “Adopting differentiated instruction to create inclusive classrooms”. Currently she is affiliated as a guest professor at the faculty of psychology and pedagogy of Ghent University. In addition she works as a professionalisation coach for Schoolmakers. Within her research Esther focuses on differentiated instruction, diversity in education, inclusive education and teacher education.

Júlia Grifol-Freixenet obtained her doctoral degree in 2020 at the Department of Educational Sciences at the Vrije Universiteit Brussel, with a dissertation titled “Learning about inclusive education: Exploring the entanglement between Universal Design for Learning and Differentiated Instruction”. Currently, Júlia works at the University of Barcelona.

Prof. Dr. Katrien Struyven is an associate professor at Hasselt University (UHasselt) and at Vrije Universiteit Brussel (VUB). Creating inclusive learning environments which address student diversity in a positive way are key to her work (differentiated instruction, PAL, assessment for learning, feedback). Katrien mainly teaches instructional science courses to students in the Teacher Training Program (UHasselt).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 31

Evaluating Effective Differentiated Instruction in Multicultural South African Secondary Schools



Thelma de Jager

Abstract Various contextual factors contribute to teachers' failure to employ effective differentiated instruction in classroom practices. This is evident in the high drop-out rate of South African students. These students mostly attend schools in lower income and rural areas that are poorly resourced and where teachers are not always trained to support students and apply differentiated instruction in the classroom. The study aimed to establish from students' observations and teachers' perceptions to what extent differentiated instruction is employed in classrooms. Data was collected in public secondary schools ($n = 25$) of the Gauteng Province in South Africa, using a quantitative approach. The social context of these schools still embodies poverty, lack of educational opportunities and resources, and overcrowded classrooms (ratio 1:40). Two questionnaires were completed, one by secondary school students ($n = 4510$) and another by teachers ($n = 424$). Contradictions were detected when students' observations of their teachers' differentiated classroom practices were compared with their teachers' perceptions. Findings showed that teachers did not always establish if students understood the content and did not always know what their difficulties were. The possible reasons could be inadequate training of teachers to identify students' learning barriers and to create and implement differentiated activities; teachers experiencing a lack of time to complete the curriculum; a lack of resources; teaching large classes; and an inability to manage and maintain discipline in classes.

Keywords Differentiated instruction · Multi-cultural · Inclusive education · Strategies and methods · Learning barriers

T. de Jager (✉)

The Department of Educational Foundation, Tshwane University of Technology,
Pretoria, South Africa

e-mail: DeJagerT@tut.ac.za

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_31

691

1 Introduction

Globally, *researchers are seeking solutions for how multicultural students* with multiple levels of academic readiness, socio-economic status, languages, intelligences, values, religions, parent education and competences and skills can be instructed effectively (Williams et al., 2009). In South Africa, a developing country, rural schools in lower income communities are mostly multicultural, poorly resourced and characterised by ineffective teaching and learning. To address these barriers to learning and linguistic and ethnic diversity in the education system, South Africa has implemented an inclusive education policy where all people are regarded equal and provided with the same opportunities and experiences to acquire effective education (Badat & Sayed, 2014). However, the inclusive education policies do not always address students' needs in effective education. Students (90.4%) are instructed in English as a second language (Fleisch, 2008); advantaged students who attend fee paying schools (mostly funded by parents) perform better than students from rural areas (funded by the government), who tend to drop out from school before completing Grade 12; and teachers are generally not adequately trained and equipped to apply differentiated instruction to address their students' learning barriers (Landsberg et al., 2011). Chataika et al. (2012) point out that teachers who lack skills in how to identify students' barriers and adjust their teaching according to diverse student needs in their classrooms impede academic progress. This is evident in the high dropout rate (60%) of South African students before completing Grade 12 (Hartnack, 2017). To address the high dropout rate, it is important that teachers should be trained and equipped with the necessary skills to create and apply differentiated instruction to accommodate diverse and individual student needs in poorly resourced schools (Brand et al., 2012).

Teachers can be considered as the main source of effective teaching and learning and application of differentiated instruction (Coe et al., 2014). Therefore, it is important to determine the extent to which secondary school teachers support students and apply differentiated instruction in their classrooms to address the needs of the increasingly multicultural body of students, particularly in South Africa. Additionally, students' experiences and observations of teachers' support and application of differentiated teaching practices could add value in how their needs can be addressed.

2 Literature Review

The united South African population comprises diverse religions and cultures. These multicultural groups form part of the country's heritage, identity and culture, where the goal is to help each culture understand and respect other cultural practices and to unite all South African citizens.

Before 1993, South African education was characterised by an apartheid system in which students attended separate schools according to race (Msila, 2007). During this period, 'Black' schools were characterised by ineffective education; overcrowded classrooms; teacher-centred instruction; under- and unqualified teachers; inadequate resources; reduced school attendance rates of students and teachers; conflict, violence and disruptions in and around schools; and poor academic achievement. Mother tongue instruction had been the norm in African schools for the first eight years of schooling (Centre for Development and Enterprise [CDE], 2015). At that stage the majority of students wanted to be instructed in English rather than their mother tongue, unaware of the potential benefits of mother tongue development at an early age (Higgs & Van Wyk, 2007).

The post-apartheid education policies established a single education system for all national cultures, new education managers were appointed, and curricula revised (Lekgoathi, 2010). Despite these radical changes and curriculum revisions, in 2003 South Africa scored the lowest of 50 countries in the *Trends in International Mathematics and Science Study* (TIMSS) that tested Grade 8 mathematics and science proficiency of students (Spaull, 2013). The Department of Basic Education (2013) realised that effective education commences in early childhood education, where students are instructed in ESL and not their home language. Therefore, the Annual National Assessments (ANA) were implemented in 2014 to test students' language and numeracy skills (Department of Basic Education, 2018). The ANA tests, managed by the schools themselves, include standardised Home Language, First Additional Language and Mathematics tests and are written by all students in Grades 1 to 6 and 9. The 2013 results showed the following average percentage marks: Home Language = 44.0%, Second Home Language = 38.1% and Mathematics = 15.9% (Department of Basic Education, 2013). The tests indicated that mother tongue instruction could contribute to students' effective learning.

Furthermore, more than two decades ago, McAdamis (2001) established that poor performing students' test scores could be improved when employing differentiated instruction. The study also noticed that students experiencing such differentiated approaches showed more enthusiasm and motivation to learn. Equally important, a study in Iran on female students using differentiated instruction to teach vocabulary in mixed-ability classes showed a positive impact on the students' academic performance (Alavinia & Farhady, 2012). Moreover, a study conducted in Kenya by Muthomi and Mbugua (2014) found that differentiated instruction improved secondary school students' achievement in mathematics significantly.

Since 1991, Bourdieu and Coleman propagated that the economic, social, cultural, and political values of the country in large part determine effective education. In South Africa, the number of students excluded from the education system, socio-economic status, and availability of support structures differ from province to province, among school systems, and from community to community (UNESCO, 2003). In addition, to apply differentiated instruction successfully in multicultural schools of South Africa, it is important to understand the constitution of the educational system.

The educational system of South Africa consists of two types of schools namely independent and public schools. Public schools are state controlled and independent schools are privately governed. Most of the students attend public schools ($n = 23,796$), while a minority of students attend independent schools ($n = 1966$) due to high school fees (Statistics South Africa, 2019).

2.1 Importance of Differentiated Instruction in Effective Teaching

Various contextual factors contribute to the substandard quality of teaching in South Africa, such as frequent power outages; absenteeism of teachers; ill-equipped and large classes; teachers (12%) diagnosed with AIDS/HIV; lack of teaching and learning resources; students with insufficient reading and writing skills; multi-cultures; poverty; poor school management and leadership; lack of parental involvement in their children's education; students' linguistic and cultural diversity; sexual abuse of student girls – often by male school teachers; pregnancy; and inadequately trained teachers who are not always able to adapt their teaching methods and strategies effectively to students' needs (Bernstein, 2015; Spaul, 2013).

The increasing diversity of classrooms and the inclusion of multicultural students with different learning abilities demand culturally sensitive and differentiated instruction that provides for the development of the whole individual (Anderson, 2007). In addition, Mpofo et al. (2014) emphasised the alignment of content with local cultures that includes values, beliefs, experiences, behaviours, and other characteristics of diverse cultures in achieving effective teaching and learning. The connection of new content to students' prior learning which derives from real-life experiences is not only viewed as a cultural border crossing but also a crutch to understand new content (De Jager, 2019).

Differentiated instruction can be constructed from various theories, such as instruction responsive to students' various interests; depiction of the readiness levels and learning profiles of students (Tomlinson, 2005); adjustment of the learning environment content, process and product for effective learning (Rock et al., 2008); supportive and adjustable teaching materials, methods and strategies that teachers use to include all students in the learning activities regardless of their differences in ability (De Jager, 2013, 2017); various ways to include different learning preferences and students' individual interests (Anderson, 2007); and understanding how students assimilate and understand facts (Anderson, 2007). Thus, differentiation can be described as flexible but organised ways of proactively adjusting teaching and learning methods to accommodate students' various learning preferences and needs in achieving maximum growth and development to reach their full potential.

Contrary to traditional teacher and textbook-centred learning methods, differentiated learning activities are student-centred, where the students are responsible for their own learning. Differentiated teaching allows students to engage in individualised activities and collaborative discussions among their peers. Thus, students could acquire extra assistance from their peers to solve a problem rather than using only

the teacher as the sole instructor. In agreement, research of Payne et al. (2004) shows that group work assists students to engage actively, develop teamwork skills and learn new content more in depth from one another. Elaborating, Genzuk (2011) points out that teachers could apply diverse student-centred teaching strategies, which include direct instruction, hands-on activities and visual aids, to connect new content to prior knowledge that could assist students to understand the new content, and allow students to process new information at their own pace.. These strategies could allow students to process meaning to new and abstract concepts while learning at their own readiness level.

However, the application of differentiated instruction is often hindered by: *teachers' unwillingness to create differentiated activities due to a heavy workload, insufficient resources, pressure to complete a large amount of content in a limited time, teach large classes and lack sufficient training in differentiated teaching practices* (Dalton et al., 2012). This results in teachers employing mostly teacher centered "talk and chalk" methods which could contribute to poor academic performance of South African students.

Moreover, Spaul (2013) points out that even though the South African education policy requires education circuit and district offices to observe, evaluate and support teachers' teaching practices, these evaluations seldom occur. In search of a solution, Ampadu (2012) suggests that students' views of teachers' teaching practices could enhance effective learning as students could become more engaged in active learning when they experience that their voice is important. Wallace et al. (2016) agree that students' perceptions of how they learn during classroom interactions are essential for effective education. In addition, Bourke and Mentis (2013) emphasise that the acknowledgement of students' perceptions can contribute to a significant development and improvement of differentiated instruction. On the other hand, Rantanen (2012) warns that students might use the opportunity to evaluate the teacher on a more personal level, which could be biased. Göllner et al. (2018) point out that students can observe the same teacher's classroom practices differently and could be influenced by personal preferences according to a teacher's popularity or the manner in which they address their individual needs.

It is also found that teachers' ratings of their classroom practices and their students' perceptions about actual differentiated teaching practices might differ (Kunter & Baumert, 2006). Thus, two different perspectives which include students' views as active participants in the classroom and perceptions of their teachers on the employment of diverse teaching approaches to support students, could add significantly to the development of differentiated instruction.

2.2 Aims of the Study

The aims of the study were to detect from perceptions of secondary school teachers whether they support students in applying differentiated instruction in teaching practices and to establish from students' views whether their teachers were applying differentiated instruction.

Students' perceptions were integrated in the study, based on research findings of Ampadu (2012), Anderson and Miller (1997), and Bansilal et al. (2010), who found that students' evaluations of their teachers' teaching practices proved to be reliable and viable. This is because the application of instruction methods has a significant impact on students' learning experiences. Moreover, Feistauer and Richter (2017) indicate that very few studies are available on students' perceptions of their teachers' teaching practices.

Therefore, the first aim of the current study was to determine from students' experiences how effectively their teachers employ differentiated instruction to address their learning needs. The problem is encapsulated in the following research question:

- According to students' experiences, how effectively are teachers applying differentiated instruction in secondary school classes?

The second aim of the study was to determine from secondary school teachers' perceptions the extent they utilised differentiated instruction in poorly resourced schools and support their students. More specifically, this study sought to provide answers to the following question:

- How do secondary school teachers apply differentiated instruction in their teaching practices to address students' learning needs?

3 Methods

3.1 Procedure and Sample

Quantitative data was collected in randomly selected public secondary schools ($n = 25$) of the Gauteng Province in South Africa. The research included secondary school students ($n = 4510$) of diverse cultures and their teachers ($n = 424$), who all voluntarily agreed to participate in this study.

The social context of these randomly selected public secondary schools still embodies poverty, a lack of educational opportunities and resources, and overcrowded classrooms (ratio 1:40). The Gauteng Province was selected because it hosts more than 25% (14 million) of the population, although it is the smallest of nine provinces, has the highest secondary school completion rate (72%) followed by the Western Cape Province (70%), and is responsible for a third of South Africa's income (Statistics South Africa, 2016). In addition, the Grade 12 final examination results of the Gauteng Province do not deviate significantly from other provinces.

3.2 Research Design

A quantitative approach was used, firstly, to determine secondary school students' experiences of their teachers' differentiated classroom practices and secondly, to establish from teachers' perceptions to what extent they employed differentiated instruction in their classes. Questionnaires were completed by secondary school students in how they observed their teachers' differentiated teaching practices in class. Additionally, another questionnaire was used where teachers could voice their perceptions on their teaching practices and establish to what extent they are applying differentiated instruction to support students. Questions and answers related to differentiated instruction were purposively sampled from the ICALT3 questionnaires in finding answers to the research questions of this study. The ICALT 3 questionnaires were compiled from research studies by Danielson (2013), Pianta and Hamre (2009) and Van de Grift (2007) and tested in countries experiencing similar education challenges as South Africa (e.g. the Slovak Republic, very rural parts of Scotland, and Croatia) (Maulana et al., 2014). Previous research findings deriving from the ICALT3 questionnaires indicate the reliability and validity of the measuring instrument applied in this study..

The sampled questions (from ICALT3 student and teacher questionnaires) related to applied differentiated instruction in lessons. The aim of using two sets of questionnaires was important to detect in depth to what extent South African teachers apply differentiated instruction in the socio-context they are teaching gathered from Students' and teachers' views.

3.3 Procedures

Permission to conduct the research was obtained from the Gauteng Department of Basic Education, the school principals, teachers, and the parents of participating students. Students were requested to indicate the extent that their teachers' employ differentiated teaching practice and teachers their perceptions when they completed the ICALT 3 questionnaires. An average of 10–15 students of each participating teacher completed the questionnaire.

The data was collected over a three-month period and the anonymity of all participants was respected. The questionnaires were completed on an optical mark recognition (OMR) form. After completing the questionnaires sampled questions (See Tables 31.1 and 31.2) relevant to differentiated instruction were analysed and discussed.

Participants completed the questionnaires using a four-point Likert scale with the options 'never', 'seldom', 'frequently' and 'often'. The responses were further grouped in 'agree' ('frequently' and 'often') or 'disagree' ('never', 'seldom') to

Table 31.1 Students' perceptions of differentiated instruction in classes

	Disagree (never/ seldom)	Agree (frequently/ often)	Missing
<i>My teacher takes into account what I already know.</i>	(n = 1614) 35.8%	(n = 2732) 60.6%	(n = 164) 3.6%
<i>My teacher makes connections to what I already know.</i>	(n = 1530) 34%	(n = 2934) 65.1%	(n = 46) 1%
<i>My teacher checks whether I have understood the content of the lesson.</i>	(n = 3060) 67.8%	(n = 1340) 29.7%	(n = 110) 2.4%
<i>My teacher knows what I have difficulty with.</i>	(n = 2745) 61%	(n = 1643) 36.4%	(n = 122) 2.7%

Table 31.2 Teachers' perceptions of applied differentiated instruction in their classes

Secondary school teachers (n = 262)	Disagree	Agree
<i>When the students of this class do not perform according to their abilities, I make time to support them with extra help.</i>	9% (n = 24)	91% (n = 248)
<i>Sometimes I feel I cannot assist all students when they need me.</i>	35% (n = 92)	65% (n = 170)
<i>When the students of this class do not comprehend the lesson material, I use another approach.</i>	5% (n = 13)	95% (n = 249)
<i>If students do not understand the content of the lesson, I explain it in different ways.</i>	3% (n = 8)	97% (n = 254)
<i>I cannot tell if students are keeping up with me.</i>	29% (n = 76)	71% (n = 186)
<i>I show students different ways of how to solve a problem</i>	42% (n = 110)	58% (n = 152)
<i>I create various learning activities that students can choose from.</i>	37% (n = 97)	63% (n = 165)
<i>I have to guide students step by step when executing the activities.</i>	21% (n = 55)	79% (n = 206)
<i>I cannot allow students in this class to work on their own.</i>	44% (n = 115)	56% (n = 147)

assist the interpretation of findings. After completion of the OMR questionnaires, the data was electronically scanned and analysed. In the study descriptive research was used to explain and interpret to what extent differentiated instruction is employed in secondary school classrooms. Data obtained from the sampled questions related to differentiated instruction, was statistically analysed using the software SPSS (Version 23.0) programme and are summarised in Tables 31.1 and 31.2.

3.4 *Students' Experiences of Teachers' Differentiated Classroom Practices*

The study revealed that most of the students (56%) had been taught for at least one year by a teacher involved in this study, followed by 24.3% taught for two years by the same teacher, while only 3.5% had been taught for 0–11 months by the observed teacher. It can be concluded that most of the participants were familiar with the differentiated teaching practices of their teachers and *were able to contribute valuable findings to this study*.

Significant low scores reflected when teachers employed differentiated activities in class. This category revealed low positive scores for several teaching practices. The responses of students show that only 60.6% of the teachers considered what students already knew and 65.1% of teachers makes connections to what they already know. One would expect that in an inclusive multicultural teaching environment where ESL students grow up in various social contexts, teachers would connect new content to students' prior background knowledge as they are instructed in a second language and do not always understand difficult concepts. 'Connecting new content to students' prior knowledge could help them to understand abstract concepts from previous experiences which could impact multicultural students' academic success. A limited number of teachers 29.7% determined whether students understood the content, and only 36.4% of their teachers know what 'I have difficulty with'. The results show that ESL students experience various impediments when learning new content, that their teachers are unaware of, or unable or unwilling to address, which could lead to ineffective learning and poor academic achievement.

The poor academic performance of South African students in the ANA tests and other international tests may be connected to the low differentiated instruction results. The reasons could possibly be attached to teachers that: teach large class sizes, do not know their students' needs, do not have sufficient time to establish if all students understood the content and could be afraid of possible disciplinary problems that may occur when engaging with specific students to establish if they understood the content of the lesson.

In agreement with Landsberg et al. (2011), it can be concluded that most teachers are not effectively trained and equipped on how to include multicultural students' needs using differentiated instruction. Moreover, the results could also be linked to teachers' inability to address students' individual needs and not necessary in how teachers adapt their teaching practices in general for the whole class.

3.5 *Teachers' Perceptions of Applied Differentiated Instruction in their Classrooms*

The participating secondary school teachers (male: 49.5% and female: 50.5%) showed diverse teaching experience ranging from less than five years (21%) to above 30 years (5%). Most of them (45%) taught science subjects (i.e. Mathematics, Physical Sciences and Life Sciences), and the remaining 55% included non-science subjects (i.e. Accounting, Business Studies, Computer Application, Economics, Geography, Language, Life Orientation and Management Sciences).

Respondents 'agreed' that they: explained content in different ways (97%); used various approaches so that students could understand content (95%); 'make time to support them with extra help' (91%); 'have to guide students step by step when executing the activities' (79%); 'cannot tell if students are keeping up with the me' (71%); 'sometimes I feel I cannot assist all students when they need me' (65%); 'create various learning activities that students can choose from' (63%); 'show students different ways of how to solve a problem' (58%); 'cannot allow students in this class to work on their own' (56%).

It can be concluded that the teacher participants were unsure whether they could allow their students to work on their own in class activities (44% 'agreed' and 56% 'disagreed'), some teachers showed their students alternative ways of how to solve a problem (58% 'agreed'), while others did not (42% 'disagreed'); most teachers agreed that they could not assist all students when they needed them (65%), while (35%) felt that they could.

4 Key Findings

Students' experiences of their teachers' differentiated teaching practices are important for improving effective teaching and learning (Ampadu, 2012). Although students are not trained in how to teach effectively, their observations (if not biased) contributed to valuable information in this study.

Interesting discrepancies were detected when students' observations of their teachers' classroom practices were compared with their teachers' perceptions on the employment of differentiated instruction in classes. Teachers (91%) indicated that they made time to support their students with extra help, 95% reflected that they used another approach if students did not comprehend the lesson and 97% of the teachers agreed that they explained in different ways if students did not understand the content. However, contradicting teachers' perceptions, only 29.7% students indicated that their teachers 'check whether they understood the content of the lesson', most students (61%) felt that their teachers did not know what they have 'difficulty with' and 67.8% do not 'check whether I have understood the content of the lesson'. Although teachers (97%) feel they sufficiently explain content in diverse ways to students they might not be able to establish if all students have grasped the

content, due to large and overcrowded classes and a curriculum that needs to be completed in a limited timeframe (CDE, 2015). This finding is in agreement with previous studies who found that teachers are not always able to identify their learners' barriers and do not know whether their students understand the new content. This is confirmed by the teachers' (71%) responses, which indicated that they could not tell whether students were keeping up with them.

In addition, 60.6% students agreed that teachers take in account what they already know and 65.1% agreed that teachers make connections to what they already know. The importance of connecting new content to students' prior learning which they obtained from real-life experiences is important to make connections to new content and understand concepts from a multicultural perspective (De Jager, 2019; Mpfu et al., 2014). Connections to prior knowledge could assist students to understand abstract concepts and contribute to effective differentiated instruction.

Moreover, 65% of the teachers felt they could not assist all students when they needed them, this is in align with students' observations that showed only 36.4% teachers actually know what difficulties they experience in class. On the other hand, Schwab et al. (2018) claim that students often tend to rate their teachers according to their ability to address their personal and individual needs and not for diverse teaching methods and strategies they employ in class to assist them with learning difficulties.

Responses of teachers (56%) showed that not all supplemented their lessons with group work. This could lead to not all students to engage actively in classes and learn new content more in depth (Payne et al., 2004). The reason could be that overcrowded classes could cause disciplinary problems and teachers want to avoid this. The other challenge could be that teachers teach large classes but not all students do have a seating place due to a lack of infrastructure in poorly resourced public schools.

To engage students actively in class, Genzuk (2011) suggests that teachers use various student-centred teaching strategies, which include self-regulated learning and explicit and implicit direct instruction, such as visual aids and hands-on activities, to connect meaning to content, and allow sufficient time for students to process new information at their own pace.

The results and previous studies indicate that education requires an intensive in-service training programme for teachers (Nel & Müller, 2010). These training programmes need to be in a specialised pedagogy such as differentiated instruction to support and improve students' academic learning. Thus, a continuous professional development programme which includes feedback from students' evaluations is essential for equipping teachers on how to apply differentiated instruction in improving their instruction strategies.

In addition, a solution for the effective multicultural teaching of ESL students (without lowering standards and students' expectations) could be for teachers to employ differentiated instruction, adapt teaching and assessment methods, and allow students to work interactively at their own pace according to their various learning preferences in achieving the lesson objectives.

5 Conclusion

Contradicting observations and perceptions show that teachers are not always aware of what students' needs are in the classroom and what their challenges are in effective learning. The possible explanation could be inadequate training of teachers to identify students' learning barriers and to create and implement differentiated activities or students evaluating teachers according to their popularity and ability to address their individual needs rather than evaluating them on classroom practices. In addition, teachers encounter various impediments that prevent them from applying differentiated instruction. These could include teaching ESL students, a lack of time to complete the curriculum, lack of resources, large classes, and an inability to manage and maintain discipline in class. Although the creation of differentiated activities may be time consuming, as with any instructional practice, fluency comes with experience. The author believes that if time and effective training were devoted to the creation of differentiated activities, less time would eventually be devoted to repeating content resulting from non-differentiated instruction. Additionally, education districts and circuit offices need to evaluate and support teachers' teaching practices (Spaull, 2013).

A follow-up study is important to establish whether in-service training workshops for teachers in public schools could improve the implementation of differentiated instruction despite the challenges they experience. Since the responses of this study represent only public secondary schools of South Africa, it is recommended that a follow-up study using the same teachers' classes should be conducted and the findings of the two studies compared to eliminate possible biased evaluations and enhance the validity of students' evaluations and teachers' perceptions.

This study shows some limitations. Besides for students' observations and teachers' perceptions on related differentiated instruction questions (sampled from the ICALT3 questionnaire), student achievement was not measured. Standardised tests to establish students' effective learning could add value to this study in establishing teachers' effective differentiated teaching practices. Additionally, this study was executed on a voluntarily basis. The participating teachers and students were only representative of public schools in the Gauteng Province and not of public schools in other provinces. Therefore, caution should be exercised when interpreting the results to broader South African contexts.

References

- Alavinia, P., & Farhady, S. (2012). Using differentiated instruction to teach vocabulary in mixed ability classes with a focus on multiple intelligences and learning styles. *International Journal of Applied Science and Technology*, 2(4), 72–79.
- Ampadu, E. (2012). Students' perceptions of their teachers' teaching of mathematics: The case of Ghana. *International Online Journal of Educational Sciences*, 4(2), 351–358.

- Anderson, K. M. (2007). Tips for teaching differentiated instruction to include all students. *Preventing School Failure, 51*(3), 41–44.
- Anderson, K., & Miller, E. D. (1997). Gender and student evaluations of teaching. *Political Science and Politics, 30*(2), 216–219.
- Badat, S., & Sayed, Y. (2014). Post-1994 south African education. The challenge of social justice. *The Annals of the American Academy of Political and Social Science, 652*(1), 127–148.
- Bansilal, S., James, A., & Naidoo, M. (2010). Whose voice matters? LEARNERS. *South African Journal of Education, 30*(1), 153–165.
- Bernstein, A. (2015). *South Africa needs thousands more and better teachers every year*. Sang Net Pulse. Retrieved from <http://www.ngopulse.org/press-release/south-africa-needs-thousands-more-and-better-teachers-every-year>
- Bourdieu, P., & Coleman, J. (1991). *Social theory for a changing society*. Westview Press.
- Bourke, R., & Mentis, M. (2013). Self-assessment as a process for inclusion. *International Journal of Inclusive Education, 17*, 854–867. <https://doi.org/10.1080/13603116.2011.602288>
- Brand, S., Favazza, A. E., & Dalton, E. M. (2012). Universal design for learning: A blueprint for all learners. *Kappa Delta Pi Record, 48*(3), 134–139.
- Centre for Development and Enterprise (CDE). (2015). *Teachers in South Africa. Supply and demand 2013–2025. Executive summary 2015*. Retrieved from <http://www.cde.org.za/wp-content/uploads/2015/03/Final-Revised-ES-TeacherSupplyandDemand2025.pdf>
- Chataika, T., Mckenzie, J. A., Swart, E., & Lyner-Cleophas, M. (2012). Access to education in Africa: Responding to the United Nations convention on the rights of persons with disabilities. *Disability & Society, 27*(3), 385–398.
- Coe, R., Aloisi, C., Higgins, S., & Elliot Major, L. (2014). *What makes great teaching? Review of the underpinning research*. Retrieved from <http://www.suttontrust.com/wp-content/uploads/2014/10/What-Makes-Great-Teaching-REPORT.pdf>
- Dalton, E. M., Mckenzie, J. A., & Kahonde, C. (2012). The implementation of inclusive education in South Africa: Reflections arising from a workshop for teachers and therapists to introduce universal design for learning. *African Journal of Disability, 1*(1), 1–7.
- Danielson, C. (2013). *The framework for teaching. Evaluation instrument*. Retrieved from <http://www.teachscape.com/frameworkforteaching/home>
- De Jager, T. (2013). Guidelines to assist the implementation of differentiated learning activities in South African secondary schools. *International Journal of Inclusive Education, 17*(1), 80–94.
- De Jager, T. (2017). Perspectives of teachers on differentiated instruction in multicultural South African secondary schools. *Studies in Educational Evaluation, 53*, 115–127.
- De Jager, T. (2019). Millennial science student teachers' views on decolonisation and culturally responsive teaching. *Journal of Education (University of KwaZulu-Natal), 76*, 185–201.
- Department of Basic Education. (2013). *Curriculum assessment policy statements (CAPS)*. Retrieved from <http://www.education.gov.za/curriculum/ncsgradesr12/caps/tabid/420/default.aspx>
- Department of Basic Education. (2018). *Annual National Assessments*. [https://www.education.gov.za/Curriculum/AnnualNationalAssessments\(ANA\).aspx](https://www.education.gov.za/Curriculum/AnnualNationalAssessments(ANA).aspx)
- Feistauer, D., & Richter, T. (2017). How reliable are students' evaluations of teaching quality? A variance components approach. *Assessment & Evaluation in Higher Education, 42*(8), 1263–1279.
- Fleisch, B. (2008). *Primary education in crisis*. Juta.
- Genzuck, M. (2011). *Specially designed academic instruction in English (SDAIE) for language minority students*. University of Southern California. Retrieved from <http://www.foshaylc.org/ourpages/auto/2011/4/4/42519949/SDAIE%20for%20Language%20Minority%20Students.pdf>
- Göllner, R., Wagner, W., Eccles, J. S., & Trautwein, U. (2018). Students' idiosyncratic perceptions of teaching quality in mathematics: A result of rater tendency alone or an expression of dyadic effects between students and teachers? *The Journal of Psychology, 110*, 709–725. <https://doi.org/10.1037/edu0000236>

- Hartnack, A. (2017). *Background document and review of key South African and international literature on school dropout*. Retrieved from <https://dgmt.co.za/wp-content/uploads/2017/08/School-Dropout-Background-Paper-Final.pdf>
- Higgs, P., & Van Wyk, B. (2007). Lifelong learning revisited: An African educational discourse. *Education as Change*, 11(1), 111–123.
- Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environment Research*, 9, 231–251. <https://doi.org/10.1007/s10984-006-9015-7>
- Landsberg, E., Krüger, D., & Swart, E. (2011). *Addressing barriers to learning. A south African perspective* (2nd ed.). Van Schaik.
- Lekgoathi, S. P. (2010). The history workshop, teacher development and outcomes-based education over the past seven years. *African Studies*, 69(1), 103–123.
- Maulana, R., Helms-Lorenz, M., & Van de Grift, W. J. C. M. (2014). Development and evaluation of a questionnaire measuring pre-service teachers' teaching behaviour: A Rasch modelling approach. *School Effectiveness and School Improvement*, 26(2), 1–26.
- McAdamis, S. (2001). Teachers tailor their instruction to meet a variety of student needs. *Journal of Staff Development*, 22(2), 1–5.
- Mpofu, V., Otulaja, F. S., & Mushayiwa, E. (2014). Towards culturally relevant classroom science: A theoretical framework focusing on traditional plant healing. *Cultural Studies of Science Education*, 9(1), 221–242.
- Msila, V. (2007). From apartheid education to the revised national curriculum statement: Pedagogy for identity formation and nation building in South Africa. *Nordic Journal of African Studies*, 16(2), 146–160.
- Muthomi, M. W., & Mbugua, Z. K. (2014). Effectiveness of differentiated instruction on secondary school students' achievement in mathematics. *International Journal of Applied Science and Technology*, 4(1), 116–122.
- Nel, N., & Müller, H. (2010). The impact of teachers' limited English proficiency on English second language learners in south African schools. *South African Journal of Education*, 30(4), 635–650.
- Payne, B. K., Monl-Turner, E., Smith, D., & Sumter, D. (2004). Improving group work: Voices of students. *Education*, 126(3), 441–448.
- Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, 38(2), 109–119.
- Rantanen, P. (2012). The number of feedbacks needed for reliable evaluation. A multilevel analysis of the reliability, stability and generalisability of students' evaluation of teaching. *Assessment and Evaluation in Higher Education*, 38(2), 224–239.
- Rock, M., Gregg, M., Ellis, E., & Gable, R. A. (2008). Reach: A framework for differentiating classroom instruction. *Preventing School Failure*, 52(2), 31–47.
- Schwab, S., Sharma, U., & Loreman, T. (2018). Are we included? Secondary students' perception of inclusive climate in their schools. *Teaching and Teacher Education*, 75, 31–39. <https://doi.org/10.1016/j.tate.2018.05.016>
- Spaull, N. (2013). *South Africa's education crisis: The quality of education in South Africa 1994–2011*. Report Commissioned by CDE. Retrieved from <http://www.section27.org.za/wp-content/uploads/2013/10/Spaull-2013-CDE-report-South-Africas-Education-Crisis.pdf>
- Statistics South Africa. (2016). *Statistics of South Africa 2016*. Statistics South Africa.
- Statistics South Africa. (2019). *Statistics of South Africa 2019*. Statistics South Africa.
- Tomlinson, C. A. (2005). Grading and differentiation: Paradox or good practice? *Theory Into Practice*, 44(3), 262–269.
- UNESCO. (2003). *Overcoming exclusion through inclusive approaches. Conceptual paper*. UNESCO.
- Van de Grift, W. J. C. M. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research*, 49(2), 127–152.

- Wallace, T. L., Kelcey, B., & Ruzek, E. (2016). What can student perception surveys tell us about teaching? Empirically testing the underlying structure of the tripod student perception survey. *American Educational Research Journal*, 53, 1834–1868. <https://doi.org/10.3102/0002831216671864>
- Williams, E., Olivier, T., & Pienaar, C. (2009). A model to revitalize teacher competence. *Acta Academica*, 41(4), 195–217.

Thelma de Jager is the Assistant Dean of the Faculty Humanities. She is a NRF rated researcher, received several awards for woman researcher and lecturer of the year, conducted several keynote addresses at conferences and authored and edited textbooks such as: General Subject Didactics, Creative Arts Education The Science to Teach and Differentiated Instruction. She is currently the project leader of the South African team for the ICALT 3 project and the British Council on Inclusive Education, T4ALL project. She has a passion to improve teaching pedagogy and her studies could impact education policy that speaks to the implementation of differentiated learning.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 32

Dealing with the Complexity of Adaptive Teaching through Collaborative Teacher Professional Development



Tijmen M. Schipper, Sui Lin Goei, and Siebrich de Vries

Abstract This chapter focuses on the challenges that teachers face in today’s heterogeneous classrooms when it comes to addressing students’ educational needs. By means of a conceptual discussion about this topic, relating to recent empirical studies in this field, we discuss whether teachers’ adaptive teaching behavior could be promoted through professional development approaches – such as Lesson Study – that focus explicitly on students’ learning. Taking students’ learning as a starting point in collaborative and classroom-based professional development approaches, one could expect that teachers gain more awareness of the variety of their students’ educational needs which, in turn, may lead to teachers better addressing these needs in classroom settings. It is argued that through such a cyclical and inquiry-based way of working, teachers may start to feel more competent and able to address the learning needs of students, leading to increasingly adaptive teaching practices. However, despite promising results in the literature, there is still much debate on the evidence of how Lesson Study influences adaptive teaching behavior in favor of *all students* and how this, in turn, impacts student learning. A “local proof route” to testing the effectiveness of Lesson Study might offer suitable directions.

T. M. Schipper (✉)

Department of Business, Media, and Law, Windesheim University of Applied Sciences, Zwolle, the Netherlands

e-mail: t.schipper@windesheim.nl

S. L. Goei

Department of Human Movement and Education, Windesheim University of Applied Sciences, Zwolle, the Netherlands

LEARN! Research Institute, VU Amsterdam, Amsterdam, the Netherlands

e-mail: s.l.goei@vu.nl

S. de Vries

Professorship of Subject Matter and Vocational Pedagogy, NHL Stenden University of Applied Sciences, Leeuwarden, the Netherlands

Teacher Education Department, University of Groningen, Groningen, the Netherlands

e-mail: s.de.vries@rug.nl

© The Author(s) 2023

R. Maulana et al. (eds.), *Effective Teaching Around the World*,
https://doi.org/10.1007/978-3-031-31678-4_32

707

Keywords Adaptive teaching · Lesson study · Professional development · Collaborative inquiry

1 Introduction

Teachers in mainstream education are increasingly expected to develop their classroom practices to an increasingly diverse set of students' individual backgrounds and educational needs (Ainscow et al., 2019; Corno, 2008; Mills et al., 2014; Schleicher, 2016). On the one hand, this is a result of a trend toward more learner-centered, constructivist approaches in education, calling for teacher adaptability (Parsons et al., 2018) which is about meeting the needs of students at every level (Dosch & Zidon, 2014; Jager et al., 2021). On the other hand, this is a result of global developments in the context of inclusive and special education, fueled by the *Salamanca Statement and Framework for Action on Special Needs Education* (UNESCO, 1994). Although this was arguably “the most significant international document that has ever appeared in the field of special education” (Ainscow et al., 2019, p. 671), it urged for major reforms of mainstream schools in order to develop inclusive education systems. As such, its influence has also become increasingly apparent where a gradual trend toward more inclusive practices has been witnessed internationally which resulted in various inclusive education policies (UNESCO, 2017).

Although teachers who adapt their teaching to their students' needs can expect broad support in education and society (Schleicher, 2016), and there is evidence supporting the claim that the most effective teachers are adaptive teachers (Kyriakides et al., 2009; Parsons et al., 2018), addressing students' individual needs turns out to be highly complicated, especially in increasingly heterogeneous classrooms (Parsons et al., 2018; Suprayogi et al., 2017; Van der Lans et al., 2018). This complexity stems from the assumption that adaptive teaching requires pedagogical content knowledge, skills, diagnosis of student learning, and an adaptive mindset and competencies (Corno, 2008; Van Geel et al., 2019; Vogt & Rogalla, 2009).

Due to this complexity and the specific competencies adaptive teaching requires, teachers often feel unprepared to adjust their curriculum and instruction to meet students' individual learning needs (Dixon et al., 2014). To address this, effective teacher professional development (PD) that specifically focuses on adaptive teaching strategies and how teacher adaptability can be supported, seems essential (Parsons et al., 2018).

We commence this chapter with a theoretical discussion about adaptive teaching and related “fuzzy constructs” (Deunk et al., 2018, p. 32). Next, we provide a brief overview of what counts as effective teacher PD according to contemporary educational research literature. Subsequently, we introduce one particular form of collaborative and classroom-based teacher PD, namely Lesson Study, that has the potential to enhance teachers' adaptive teaching competencies due to its explicit focus on students' learning (Dudley, 2013), and we show how Lesson Study can

promote adaptive teaching behavior and substantiate this by recent empirical studies in different educational and national contexts. We conclude this chapter with the most important theoretical and practical implications.

2 Theoretical Framework

2.1 *Adaptive Teaching*

Various concepts are used to refer to addressing the needs of different students in classroom contexts, such as adaptive teaching, differentiated instruction, and differentiation. As a result, various researchers argue that these concepts are often overlapping labels that lack clarity and clear operationalizations (Deunk et al., 2018; Prast et al., 2018; Suprayogi et al., 2017). As such, “The lack of definition and shared terminology in research on differentiation and associated strategies could be contributing to confusion, both within and outside academia” (Graham et al., 2020, p. 31). Due to this confusion, capturing adaptive teaching behavior by systematically measuring it might also be problematic and so far has not provided “much insight into the acting and reasoning of teachers who differentiate instruction well” (Van Geel et al., 2019, p. 53).

Despite the ambiguous use of labels for addressing students’ educational needs, Corno (2008) distinguishes adaptive teaching from other related constructs by placing it in the social and dynamic context of classroom situations. This, on the one hand, requires flexible, spontaneous, and responsive interventions from teachers, and, on the other hand, requires careful lesson planning and diagnosing of students’ progress and needs. In this definition, adaptive teaching is not only concerned with actual differentiated teaching activities prior to, during and after the lesson (Smale-Jacobse et al., 2019), but also involves having an ‘adaptive mindset’ in which a teacher “views student differences as assistive, affording, and enabling for teaching as well as student learning” (Corno, 2008, p. 171). Therefore, adaptive teaching is concerned with teachers’ careful and proactive planning of the curriculum, teaching materials and learning activities, as well as how they think about and anticipate to students’ learning needs in the social context of the classroom in order to reach the desired lesson objectives (Beltramo, 2017; Corno, 2008).

2.2 *Effective Teacher Professional Development*

The literature on teacher PD is abundant and there seems to be consensus that effective forms of teacher PD consist of ongoing, active and collaborative learning of teachers that is situated in practice, focused on students’ learning, and coherent with teachers’ beliefs (e.g., Borko et al., 2010; Desimone, 2009; Desimone & Stuckey, 2014; Schleicher, 2016; Webster-Wright, 2009). This contrasts with ‘traditional’ forms of teacher PD in which teacher learning was generally seen as “an in-service

training model, where teachers are expected to learn a clearly defined body of skills through a well-specified process, often delivered in one-shot workshops or courses taught away from the school premises” (Borko et al., 2010, p. 548).

In the current view on teacher PD, which started to develop about three decades ago (Vangrieken et al., 2017), teacher learning ideally occurs through participating in professional learning communities (PLCs) in which the former characteristics of effective PD (i.e., ongoing, active, collaborative, focused on student learning, and coherent with beliefs) are embedded. Participating in PLC’s that address these effective features of teacher PD, may have a positive impact on both teaching practice and student learning (Vangrieken et al., 2017; Vescio et al., 2008). The concept of a PLC “rests on the premise of improving student learning by improving teaching practice”, situating teacher learning in their day-to-day experiences (Vescio et al., 2008, p. 82).

There is a great variety of PLCs ranging from school-wide to department-based PLCs (Valckx et al., 2020) as well as formal, member-oriented, or formative PLCs (Vangrieken et al., 2017). A specific form of a PLC that is known for its explicit focus on how students learning (Dudley, 2013), and, as such, may contribute to supporting teachers’ adaptive teaching behavior (Norwich et al., 2020), is Lesson Study. A Lesson Study-team of teachers can be seen as a PLC (Desforges, 2015), but it is also argued that Lesson Study can create a culture for a school-wide PLC (Chichibu & Kihara, 2013). For PLCs to be effective, at least two conditions need to be in place: participants in PLCs need to be supported in processing “new understandings and their implications for teaching” and the focus of participants need to be on analyzing the impact of teaching on student learning (Timperley et al., 2007). Both conditions are generally taken into account in Lesson Study.

2.3 Lesson Study

The teacher PD approach Lesson Study originated in Japan over a century ago and spread rapidly around the globe since the late 1990s after the publication of ‘the Teaching Gap’ (Stigler & Hiebert, 1999). It is now perceived as one of the world’s fastest growing forms of teacher PD (Dudley, 2015) which may be a result of the fact that Lesson Study includes many of the features that are supposed to contribute to effective teacher PD (Lewis & Perry, 2014), as mentioned above. In Lesson Study, a small team of teachers collaboratively conduct ‘inquiry cycles’ (Lewis et al., 2012) of studying, designing, teaching, observing, and evaluating research lessons (Dudley, 2013). A research lesson is an actual classroom lesson which is generally designed to study and improve the teaching of a particular subject topic by focusing on student learning, (Lewis et al., 2012), but may also be focused on other aspects such as behavioral support (Nilvius, 2020).

At a glance, Lesson Study is a “deceptively simple” form of teacher PD (Dudley, 2015, p. 5) and has been manifested in various variations suiting different cultural contexts (Stigler & Hiebert, 2016). Despite these cultural variations, the core

elements ('big ideas') of Lesson Study entail that teachers (1) collaboratively perform research on their lessons, (2) combine practical knowledge and external knowledge, (3) learn from students, (4) make a collaborative effort through engaging in intensive professional dialogue, and (5) follow repeated cycles of research lessons (Goei et al., 2021b).

More specifically, a Lesson Study cycle consists of defining a clear research purpose, studying the curriculum and classroom material, planning the research lesson in detail, teaching the research lesson by one teacher while the other members of the Lesson Study team observe the research lesson and collect (pre-defined) data, evaluating the research lesson in a post-lesson discussion based on student data, ideally guided by a facilitator or 'knowledgeable other' (Takahashi & McDougal, 2016), and reflecting on the learning experiences (Lewis et al., 2006).

A widely-used extra dimension to Lesson Study, embedded in the UK Lesson Study model (Dudley, 2013), is the application of 'case pupils' who represent certain learner groups (attainment groupings) in the classroom. All Lesson Study phases are organized around these 'case students'. In the UK model, revising and re-teaching the research lesson are also essential parts of the Lesson Study cycle. The Dutch Lesson Study model (De Vries et al., 2016) draws on the UK variant. In this model, the Lesson Study facilitator has a pivotal role and the model "allows more room for selecting 'case students' based on behavior or other criteria" (Schipper et al., 2020b, p. 353), in addition to solely learning aspects. In a variant of this model (Goei et al., 2021a), the three-tier prevention logic (Kratochwill et al., 2007) is used to select case students, focusing on case students from tier 1 (general provision), tier 2 (targeted provision), and tier 3 (specialized provision).

Various international review studies conclude that Lesson Study is a powerful PD approach. These reviews report studies in which it becomes clear that participating in Lesson Study influences teachers' knowledge, behavior and attitudes, and that teachers become more focused on the learning of their students, and also describe the impact on the school context (De Vries et al., 2017; Huang & Shimizu, 2016; Xu & Pedder, 2015). However, these reviews mainly draw on small-scale qualitative studies and only a few large effect studies are available in this context. The What Works Clearinghouse (WWC) body in the United States, found that, out of 643 PD studies related to K-12 Mathematics education in the US, only two studies met their evidence standards and reported significant positive effects on student math proficiency (Gersten et al., 2014), of which one reported a randomized controlled trial experiment in the context of Lesson Study (Lewis & Perry, 2017). A similar effect study on Lesson Study in the United Kingdom, conducted by the Education Endowment Foundation, did not report positive effects of participating in Lesson Study on students' mathematic and reading attainment on Key Stage 2 level (Murphy et al., 2017). However, this evaluation study did show that teachers felt that Lesson Study was a powerful PD approach and reported changes to their teaching practices. Moreover, the authors stated that "There is evidence that some control schools implemented similar approaches to Lesson Study, such as teacher observation. This trial might, therefore, underestimate the impact of Lesson Study when introduced in schools with no similar activity" (p. 4).

Working with RCTs is in line with thinking about instructional improvement via the so-called general proof route (Lewis et al., 2006), while Lesson Study and working with it are more in line with the “local proof route, whereby locally initiated innovations can contribute to broad instructional improvement, with education researchers supporting the explication, development, and testing of such innovations” (Lewis et al., 2006, p. 10). In addition, we actually do not know enough yet about the nature and mechanisms of Lesson Study to test it summatively. Hence, “Controlled experimental research on immature versions of lesson study could lead us to conclude that it doesn’t work, and to move on to the next promising idea” (Lewis et al., 2006, p. 10). Moreover, other Lesson Study researchers argue for Lesson Study “to be treated holistically as a vehicle for development and improvement at classroom, school and system levels rather than as a curricular or pedagogical intervention” (Dudley et al., 2019, p. 202), and should therefore contain indicators of impact at both school and local system levels (Dudley et al., 2019).

3 Promoting Adaptive Teaching Through Lesson Study: What Do We Know?

3.1 Overview of the International Literature

Despite the growing knowledge base around Lesson Study, studies that focus on the role of Lesson Study in inclusive mainstream classroom settings, specifically addressing how the needs of all students could be addressed, are scarce. In this chapter, we present an overview of the international literature about Lesson Study in relation to adaptive teaching by clustering these studies around the contexts in which they took place. We start this chapter by presenting the studies conducted in primary education situated in different cultural, though European, contexts (Sect. 3.1.1). In the subsequent section (Sect. 3.1.2), we address the secondary education context. As we found that these studies are, so far, predominantly situated in the Dutch context, we refer to this section as ‘The Dutch case’. We conclude this chapter with a short section about Lesson Study in the special needs contexts in which focusing on students’ individual needs is generally more self-evident (Sect. 3.1.3).

3.1.1 Adaptive Teaching Through Lesson Study in Primary Education

We found four recently published studies in primary education with the topic of adaptive teaching in the context of Lesson Study research. In the Swedish context, two studies draw attention as they are specifically concerned with catering for all students and how Lesson Study could promote this. Nilvius (2020) described how the multi-tiered Response To Intervention model (RTI) can be used in Lesson Study to maximize the achievement of all students. In this pilot study, teachers claimed “that the RTI model gave them good control over all the students’ development in

basic skills and that monitoring all students' development was important to better understand their needs" (Nilvius, 2020, p. 284). A second study (Lundbäck & Egerhag, 2020), also in Swedish primary education, described how Lesson Study enhanced the mathematical learning of all students in two learning situations, including students with special needs.

In another Scandinavian country, Norway, Aas (2020) presents findings of a study in primary education where Lesson Study was used to examine teacher talk focusing specifically on inclusive and adaptive education for all students. The study shows how teachers talk about students' needs (in terms of academic needs, behavioral needs, and the learning environment) and what kind of beliefs they have about these needs. As a result of participating in Lesson Study, teachers in this study reported to have become more aware of students' needs and gained increased trust in students' abilities as well as trust in their own ability to influence students' learning and development. Moreover, the study shows how teachers changed their classroom behavior in more inclusive ways.

In the Austrian context, Mewald and Mürwald-Scheifinger (2019) describe a train-the-trainer program that emphasizes the role of knowledgeable others, established to support implementing "educational change and further competence-oriented learning" (p. 219) in primary education. Their Lesson Study program was based on the "combination of a typical lesson study cycle with six design principles" including the principle to help teachers in "providing appropriate, relevant and adaptive learning experiences aligned with their students' interests, dispositions and needs" (p. 220). In presenting the experiences of knowledgeable others in this program, one of them described that this program "changed our attitude towards pupils' learning" (p. 227). In addition, teachers reported a focus on including all students and make particular reference to students from a migrant background stating that "It was very exciting to discover that using a lesson study approach created a much greater learning growth in children with migrant backgrounds compared to those without. This finding led us to critically examine our lesson planning to find out if we are really reaching all or as many children as possible" (p. 227).

In sum, these studies in the context of primary education show that Lesson Study can impact teachers' adaptive mindset and knowledge, and this leads to differences in teachers' adaptive behavior. In the last case there is even evidence of changes in student learning. However, these studies rely predominantly on qualitative evidence and more evidence is needed from "repeated cycles that test key design features and create "actionable artifacts" to leverage learning at new sites" (Lewis et al., 2006, p.10).

3.1.2 Adaptive Teaching Through Lesson Study in Secondary Education: The Dutch Case

Following our literature search on studies about adaptive teaching through Lesson Study in the secondary education context, we only came across several studies that were conducted in the Netherlands. Moreover, these studies were closely related to

each other as they were part of the same overarching research project. Prior to presenting the findings of the studies conducted in the Netherlands, we start with providing a description of the Dutch educational context in order to better understand and interpret the findings.

Secondary schools in the Netherlands have a relatively high degree of autonomy, no national curriculum, and a highly ‘tracked’ educational system in which students are divided over various cognitive tracks based on their standardized test scores in the last grade of primary education (OECD, 2016). These tracks include practical training, pre-vocational secondary education, senior general secondary education, and pre-university education. Despite the merits of this tracked system and the opportunities to move easily from one track to another, “Tracked systems tend to deprive low-performing students of the positive peer effects from stronger students” (OECD, 2016, p. 64). In line with the earlier described trend toward more inclusive practices, the Netherlands also aims to promote inclusive policies and classroom practices through, for example, the introduction of the *Appropriate Education Act* in 2014. This act obliges school leaders in collaboration with regional partners (other schools, including special education schools) to make sure that every child is offered appropriate education suited to his or her capabilities (OECD, 2016). Despite these introduced policies, teachers in the Netherlands struggle to assess and address the increasingly diverse needs of students (Dutch Inspectorate of Education, 2020) and this applies in particular to teachers who are new to the profession (OECD, 2016). The Dutch Inspectorate of Education (2020) concludes that, despite initiatives to promote adaptive teaching through the appropriate Education Act, not all schools feel the collective responsibility in their regional partnerships to cater for all students, which may have severe consequences for individual students. Following this context description, it is not surprising that effective teacher PD, particularly focused on adaptive teaching skills, is an increasingly important way of preparing teachers to address their students’ needs (OECD, 2016). Hence, Lesson Study receives increasing attention in the Netherlands (De Vries et al., 2016), not only in the context of inclusive education.

In the presented studies below, the Lesson Study model was used in which case students were selected on the basis of the three-tier prevention logic (Goei et al., 2021a). Depending on the research theme and research questions of the Lesson Study teams – which could vary from a more content-specific focus to a more generic focus on, for example, students’ motivation – teachers studied classroom and student material and then designed the research lesson with an explicit focus on the selected case students. Subsequently, the research lesson was taught by one of the teachers and observed by the other members of the Lesson Study team, again focusing specifically on the case students’ behaviors using self-constructed observation forms. The research lesson was then discussed and evaluated based on the collected observation data and case student interviews which took place directly after the research lesson. Finally, the research lesson was revised and re-taught followed by a reflection on the complete Lesson Study cycle.

In a first qualitative and explorative study, Schipper and colleagues (2017) examined to what extent participation in at least two Lesson Study cycles during one

academic year enhanced teachers' adaptive teaching competence in terms of their knowledge, beliefs and attitudes about students' educational needs, and how teachers addressed (or tried to address) these needs in daily practice as a result of LS. This study also examined the role of the school context in promoting or hindering this. The results show that teachers gave clear notions of how Lesson Study participation increased their awareness of their students' needs and how their beliefs and attitudes about adaptive teaching changed. Teachers also reported either incidental or structural changes in their adaptive teaching behavior. What contributed most to these changes were an explicit focus on student learning in Lesson Study, the ample opportunities in Lesson Study that allow to experiment with adaptive teaching strategies, and the guiding role of the Lesson Study facilitator. In terms of the school context, support of the school leader, learning from colleagues, and sufficient time were found essential in promoting these practices.

In a second study conducted by the same authors (Schipper et al., 2018), a quasi-experimental mixed-methods design was used to examine the influence of participating in Lesson Study on teachers' adaptive teaching behavior. As teacher self-efficacy, defined as "teachers' belief or conviction that they can influence how well students learn, even those who may be difficult or unmotivated" (Guskey & Passaro, 1994, p. 628), was related to more positive attitudes toward adaptive teaching practices (Suprayogi et al., 2017), the study also addressed the influence of participating in Lesson Study on teachers' self-efficacy beliefs and the relation between adaptive teaching and teacher self-efficacy. The results showed a significant intervention effect for the subscale 'efficacy in student engagement' and a positive within-group effect on the subscale 'efficacy in instructional strategies', indicating that teachers who participated in Lesson Study felt more capable to engage all students in their lessons and to use various strategies in their instruction. Teacher behavior was measured using the ICALT observation instrument (Van de Grift, 2007). Although intervention effects were found for the subscales 'efficient classroom management' and 'clarity of instruction' in favor of the Lesson Study group, no intervention effects were found for the adaptive teaching domain. With stimulated recall interviews, the researchers were able to learn more about teachers' thoughts and actions during their lessons. It was found that teachers who participated in Lesson Study, expressed more awareness of students' educational needs and these teachers claimed that Lesson Study allowed them to experiment with adaptive teaching strategies and material.

To determine whether the self-reported findings in the first two studies could be supported by classroom observation data, a third study by Schipper and colleagues (2020c) examined adaptive teaching in more detail, again using a quasi-experimental mixed-methods design. For the purpose of this study, an observation instrument was constructed for which the ICALT observation instrument "was used as an anchor to assess the validity" (p. 7). Although the observation instruments did not yield any significant intervention effects in terms of adaptive teaching behavior, the qualitative data showed that teachers who participated in Lesson Study indicated that Lesson Study played an important role in becoming more aware of students' needs and supported them in addressing (or trying to address) these needs accordingly.

They particularly valued the use of case students in this process. The fact that, overall, the observation instruments did not capture the growth in adaptive teaching behavior that was reported by teachers in the stimulated recall interviews was found to be remarkable. Several potential reasons for this conflicting difference in output were related to the complexity of adaptive teaching, both in terms of teachers' conceptualizations of this construct, which showed a great variety of how teachers' defined and perceived adaptive teaching, as well as how to measure this construct as observers did not have information about the students, their educational needs, students' previous experiences with the subject, and teacher-student relationships.

The studies in secondary education show how participating in Lesson Study can impact teachers' adaptive mindset and adaptive teaching competence, but the results are not conclusive as the self-report evidence is not supported by the observation data. In these studies, however, it was argued that more time would be needed to see actual changes in adaptive teacher behavior and observers would need more knowledge about teachers' decisions in terms of adaptive teaching and their teacher-student relationships. As a result, we can conclude that more evidence is needed about the actual impact of participating in Lesson Study on adaptive teaching behavior given the local context in which it takes place, and, more specifically, what mechanisms in Lesson Study influence adaptive teaching behavior.

3.1.3 Lesson Study in Special Needs Education

Based on a recent literature review about the use of Lesson Study in the context of inclusive and special needs education (Norwich et al., 2020), a recent special issue in the *International Journal for Lesson and Learning Studies* (IJLLS) about perspectives of PD in special didactics (2020, Volume 9, Issue 3), and the recently published book entitled 'Lesson Study in Inclusive Educational Settings' (Goei et al., 2021a), it becomes clear that inclusive education and special needs education receive increasing attention in Lesson Study research. Studies conducted in this context are primarily concerned with using Lesson Study as a means to enhance teachers' knowledge and skills so that they can adapt their teaching to students with special educational needs in inclusive settings. This, for example, refers to applying Lesson Study to address the needs of students with neurodevelopmental conditions (Leifler, 2020), mild-to-moderate intellectual disabilities (Klefbeck, 2020), and moderate learning difficulties (MLD) (Norwich & Ylonen, 2013). In the last case, Norwich and Ylonen (2013) followed a local proof route using a realist evaluation methodology to take contextual conditions into account and found that Lesson Study "enabled teachers to develop teaching approaches and a focus on the learning requirements of pupils with MLD, who then showed some gains in their learning" (p. 171). Students in this study were assessed using different measures on reasoning, literacy and motivation.

4 Conclusion and Discussion

Adaptive teaching receives increasing attention due to an international trend toward more inclusive practices and the notion that teacher adaptability is linked to effective teaching (Kyriakides et al., 2009; Parsons et al., 2018). This chapter presented an overview of the current literature on adaptive teaching and whether the collaborative and classroom-based PD approach Lesson Study could support teachers in the increasingly complex endeavor of adapting their behavior to their students' educational needs. Based on the available international literature, we argue that Lesson Study indeed has the potential to promote teachers' adaptive teaching behavior, but much is still unknown about its effectiveness. We believe that the local proof route can contribute to this in order to find out more about the working mechanisms in Lesson Study that impact teacher behavior and student learning in turn. We also believe that a variety of methodologies, including a cross-disciplinary and cross-cultural approach, would benefit the knowledge base around Lesson Study.

In the presented studies, conducted in different European contexts, teachers appeared to be (very) positive about the potential of Lesson Study in preparing teachers for inclusive teaching practices. In general, teachers seemed to gain more awareness of their students' educational needs and gained more knowledge and skills needed to address these needs as a result of participating in Lesson Study. Awareness was enhanced in different ways, for example by closely examining and discussing student behavior, by writing down expectations of student behavior prior to the research lesson, and by interviewing the case students (Schipper et al., 2017). This impacted the way they prepared and executed their lessons by focusing on what students actually need in order to meet the learning objectives. This is most likely the result of taking student learning as a starting point by organizing research lessons around case students (Dudley, 2013). The Response to Intervention model that was used in the Swedish (Nilvius, 2020) and Dutch context (Schipper et al., 2017), may be particularly supportive in selecting these case students and making sure that a representation of all students in the classroom are included in the Lesson Study process. Future studies in the context of Lesson Study in inclusive settings may further examine this.

Despite the added value of the various studies presented in this chapter, it also becomes clear that research on Lesson Study focusing specifically on adaptive teaching is still in its infancy. After all, studies generally focus on special needs students and tend to be situated in primary education. Therefore, clear evidence of how Lesson Study influences adaptive teaching behavior in favor of *all students* and how this, in turn, impacts student learning is still lacking. Capturing adaptive teaching behavior in the classroom using objective measures, proved to be extremely complex and we argue that this is a result of the diffuse conceptualizations of adaptive teaching and the way the concept is operationalized (Deunk et al., 2018).

Finally, in order to draw conclusions about the effectiveness of Lesson Study in terms of influencing adaptive teaching behavior, we argue that school contextual conditions should be taken into account in Lesson Study research. School leaders,

for example, play an essential role in the implementation and sustainability of Lesson Study practices in order to promote adaptive teaching behavior in schools. This essential role not only refers to providing the needed structural conditions (e.g., available time to participate in Lesson Study) and cultural conditions (e.g., a shared vision and collegial support) in the school (Schipper et al., 2020a), but also to having a thorough understanding of Lesson Study and the implications for the school structures and cultures (Seleznyov, 2019) in order for Lesson Study to become an organizational routine (Wolthuis et al., 2020). In addition, even if the school context is very supportive for implementing and sustaining Lesson Study practice, much relies on teachers' adaptive teaching competencies and their motivation, mindset and ideals when it comes to becoming more adaptive teachers. Therefore, we should "acknowledge the slow and incremental way in which teachers incorporate new ideas into their ongoing practices" (Kennedy, 2016, p. 973).

References

- Aas, H. K. (2020). Professional development for inclusive and adaptive education: Lesson Study in a Norwegian context. *Professional Development in Education*. <https://doi.org/10.1080/19415257.2020.1850509>.
- Ainscow, M., Slee, R., & Best, M. (2019). Editorial: The Salamanca statement: 25 years on. *International Journal of Inclusive Education*, 23(7–8), 671–676.
- Beltramo, J. L. (2017). Developing adaptive teaching practices through participation in cogenerative dialogues. *Teaching and Teacher Education*, 63, 326–337.
- Borko, H., Jacobs, J., & Koellner, K. (2010). Contemporary approaches to teacher professional development. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (Vol. 7, 3rd ed., pp. 548–556). Elsevier.
- Chichibu, T., & Kihara, T. (2013). How Japanese schools build a professional learning community by lesson study. *International Journal for Lesson and Learning Studies*, 2(1), 12–25.
- Corno, L. (2008). On teaching adaptively. *Educational Psychologist*, 43(3), 161–173.
- De Vries, S., Verhoef, N., & Goei, S. L. (2016). *Lesson study: een praktische gids voor het onderwijs* [Lesson Study: A practical guide for education]. Garant Publishers.
- De Vries, S., Roorda, G., & Van Veen, K. (2017). *Lesson study: Een effectieve en bruikbare professionaliseringsbenadering voor de Nederlandse context?* [Lesson Study: an effective and useful professional development approach in the Dutch context?]. University of Groningen.
- Desforges, C. (2015). Lesson study as a strategic choice for CPD. In P. Dudley (Ed.), *Lesson study. Professional learning for our time* (foreword). Routledge.
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–199.
- Desimone, L. M., & Stuckey, D. (2014). Sustaining professional development. In L. Martin, S. Kragler, D. Quatroche, & K. Bauserman (Eds.), *Handbook of professional development in education: Successful models and practices* (Prek-12) (pp. 467–482). Guilford Publications.
- Deunk, M. I., Smale-Jacobse, A. E., De Boer, H., Doolaard, S., & Bosker, R. J. (2018). Effective differentiation practices: A systematic review and meta-analysis of studies on the cognitive effects of differentiation practices in primary education. *Educational Research Review*, 24, 31–54.
- Dixon, F. A., Yssel, N., McConnell, J. M., & Hardin, T. (2014). Differentiated instruction, professional development, and teacher efficacy. *Journal for the Education of the Gifted*, 37(2), 111–127.

- Dosch, M., & Zidon, M. (2014). "The course fit us": Differentiated instruction in the college classroom. *International Journal of Teaching and Learning in Higher Education*, 26(3), 343–357.
- Dudley, P. (2013). Teacher learning in lesson study: What interaction-level discourse analysis revealed about how teachers utilised imagination, tacit knowledge of teaching and fresh evidence of pupils learning, to develop practice knowledge and so enhance their pupils' learning. *Teaching and Teacher Education*, 34, 107–121.
- Dudley, P. (2015). How lesson study works and why it creates excellent learning and teaching. In P. Dudley (Ed.), *Lesson study. Professional learning for our time* (pp. 1–28). Routledge.
- Dudley, P., Xu, H., Vermunt, J. D., & Lang, J. (2019). Empirical evidence of the impact of lesson study on students' achievement, teachers' professional learning and on institutional and system evolution. *European Journal of Education*, 54(2), 202–217.
- Dutch Inspectorate of Education. (2020). *De Staat van het Onderwijs 2020* [The State of Education 2020]. Dutch Inspectorate of Education.
- Gersten, R., Taylor, M. J., Keys, T. D., Rolffhus, E., & Newman-Gonchar, R. (2014). *Summary of research on the effectiveness of math professional development approaches* (No. REL2014-010). U.S. Department of Education, Institute of Education Sciences, National Center for Educational Evaluation and Regional Assistance, Regional Educational Laboratory Southeast.
- Goei, S. L., Norwich, B., & Dudley, P. (2021a). *Lesson study in inclusive educational settings*. Routledge.
- Goei, S. L., Van Joolingen, W. R., Goetsch, F., Khaled, A., Coenen, T., In 't Veld, S. G. J. G., De Vries, S., & Schipper, T. M. (2021b). Online lesson study: Virtual teaming in a new normal. *International Journal for Lesson & Learning Studies*. <https://doi.org/10.1108/IJLLS-09-2020-0078>
- Graham, L. J., De Bruin, K., Lassig, C., & Spandagou, I. (2020). A scoping review of 20 years of research on differentiation: investigating conceptualization, characteristics, and methods used. *Review of Education*. <https://doi.org/10.1002/rev3.3238>.
- Guskey, T. R., & Passaro, P. D. (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal*, 31(3), 627–643.
- Huang, R., & Shimizu, Y. (2016). Improving teaching, developing teachers and teacher educators, and linking theory and practice through lesson study in mathematics: An international perspective. *ZDM Mathematics Education*, 48, 393–409.
- Jager, M., Denessen, E., Cillessen, A. H. N., & Meijer, P. C. (2021). Sixty seconds about each student-studying qualitative and quantitative differences in teachers' knowledge and perceptions of their students. *Social Psychology of Education*, 24, 1–35.
- Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86(4), 945–980.
- Klefbeck, K. (2020). Lesson study for students with intellectual disability. *International Journal for Lesson and Learning Studies*, 9(3), 245–259.
- Kratochwill, T. R., Volpiansky, P., Clements, M., & Ball, C. (2007). Professional development in implementing and sustaining multitier prevention models: Implications for response to intervention. *School Psychology Review*, 36(4), 618–631.
- Kyriakides, L., Creemers, B. P. M., & Antoniou, P. (2009). Teacher behaviour and student outcomes. Suggestions for research on teacher training and professional development. *Teaching and Teacher Education*, 25, 12–23.
- Leifler, E. (2020). Teachers' capacity to create inclusive learning environments. *International Journal for Lesson and Learning Studies*, 9(3), 221–244.
- Lewis, C., & Perry, R. (2014). Lesson study with mathematical resources: A sustainable model for locally-led teacher professional learning. *Mathematics Teacher Education and Development*, 16(1), 22–42.
- Lewis, C., & Perry, R. (2017). Lesson study to scale up research-based knowledge: A randomized, controlled trial of fractions learning. *Journal for Research in Mathematics Education*, 48(3), 261–299.

- Lewis, C., Perry, R., & Murata, A. (2006). How should research contribute to instructional improvement? The case of lesson study. *Educational Researcher*, 35(3), 3–14.
- Lewis, C. C., Perry, R. R., Friedkin, S., & Roth, J. R. (2012). Improving teaching does improve teachers: Evidence from lesson study. *Journal of Teacher Education*, 63(5), 368–375.
- Lundbäck, B., & Egerhag, H. (2020). Lesson study as a bridge between two learning contexts. *International Journal for Lesson and Learning Studies*, 9(3), 289–299.
- Mewald, C., & Mürwald-Scheifinger, E. (2019). Lesson study in teacher development: A paradigm shift from a culture of receiving to a culture of acting and reflecting. *European Journal of Education*, 54(2), 218–232.
- Mills, M., Monk, S., Keddie, A., Renshaw, P., Christie, P., Geelan, D., & Gowlett, C. (2014). Differentiated learning: From policy to classroom. *Oxford Review of Education*, 40(3), 331–348.
- Murphy, R., Weinhardt, F., Wyness, G., & Rolfe, H. (2017). *Lesson study. Evaluation report and executive summary November 2017*. Education Endowment Foundation.
- Nilvius, C. (2020). Merging lesson study and response to intervention. *International Journal for Lesson and Learning Studies*, 9(3), 277–288.
- Norwich, B., & Ylonen, A. (2013). Design based research to develop the teaching of pupils with moderate learning difficulties (MLD): Evaluating lesson study in terms of pupil, teacher and school outcomes. *Teaching and Teacher Education*, 34, 162–173.
- Norwich, B., Benham-Clarke, S., & Goei, S. L. (2020). Review of research literature about the use of lesson study and lesson study-related practices relevant to the field of special needs and inclusive education. *European Journal of Special Needs Education*. <https://doi.org/10.1080/008856257.2020.1755929>
- OECD. (2016). *Netherlands 2016: Foundations for the future. Reviews of National Policies for education*. OECD Publishing.
- Parsons, S. A., Vaugh, M., Scales, R. Q., Gallagher, M. A., Parsons, A. W., Davis, S. G., Pierczynski, M., & Allen, M. (2018). Teachers' instructional adaptations: A research synthesis. *Review of Educational Research*, 88(2), 205–242.
- Prast, E. J., Van de Weijer-Bergsma, E., Kroesbergen, E. H., & Van Luit, J. E. H. (2018). Differentiated instruction in primary mathematics: Effects of teacher professional development on student achievement. *Learning and Instruction*, 54, 22–34.
- Schipper, T., Goei, S. L., De Vries, S., & Van Veen, K. (2017). Professional growth in adaptive teaching competence as a result of lesson study. *Teaching and Teacher Education*, 68, 289–303.
- Schipper, T., Goei, S. L., De Vries, S., & Van Veen, K. (2018). Developing teachers' self-efficacy and adaptive teaching behaviour through lesson study. *International Journal of Educational Research*, 88, 109–120.
- Schipper, T. M., De Vries, S., Goei, S. L., & Van Veen, K. (2020a). Promoting a professional school culture through lesson study? An examination of school culture, school conditions, and teacher self-efficacy. *Professional Development in Education*, 46(1), 112–129.
- Schipper, T. M., Goei, S. L., Van Joolingen, W. R., Willemse, T. M., & Van Geffen, E. C. (2020b). Lesson study in Dutch initial teacher education explored: Its potential and pitfalls. *International Journal for Lesson and Learning Studies*, 9(4), 351–365.
- Schipper, T. M., Van der Lans, R. M., De Vries, S., Goei, S. L., & Van Veen, K. (2020c). Becoming a more adaptive teacher through collaborating in lesson study? Examining the influence of lesson study on teachers' adaptive teaching practices in mainstream secondary education. *Teaching and Teacher Education*, 88, 102961.
- Schleicher, A. (2016). *Teaching excellence through professional learning and policy reform: Lessons from around the world* (International summit on the teacher profession). OECD Publishing.
- Seleznov, S. (2019). Lesson study beyond Japan: Evaluating impact. *International Journal for Lesson and Learning Studies*, 8(1), 2–18.

- Smale-Jacobse, A. E., Meijer, A., Helms-Lorenz, M., & Maulana, R. (2019). Differentiated instruction in secondary education: A systematic review of research evidence. *Frontiers in Psychology, 10*, 2366. <https://doi.org/10.3389/fpsyg.2019.02366>
- Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. Free Press.
- Stigler, J. W., & Hiebert, J. (2016). Lesson study, improvement, and the importing of cultural routines. *ZDM Mathematics Education, 48*, 581–587.
- Suprayogi, M. N., Valcke, M., & Godwin, R. (2017). Teachers and their implementation of differentiated instruction in classrooms. *Teaching and Teacher Education, 67*, 291–301.
- Takahashi, A., & McDougal, T. (2016). Collaborative lesson research: Maximizing the impact of lesson study. *ZDM Mathematics Education, 48*(4), 513–526.
- Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration*. Ministry of Education. <http://educationcounts.edcentre.govt.nz/goto/BES>
- UNESCO. (1994, June 7–10). *The Salamanca statement and framework for action on special needs education*. World Conference on Special Needs Education: Access and Quality. Salamanca, Spain. <https://unesdoc.unesco.org/ark:/48223/pf0000098427>
- UNESCO. (2017). *A guide for ensuring inclusion and equity in education*. UNESCO.
- Valckx, J., Vanderlinde, R., & Devos, G. (2020). Departmental PLCs in secondary schools: The importance of transformational leadership, teacher autonomy, and teachers' self-efficacy. *Educational Studies, 46*(3), 282–301.
- Van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research, 49*(2), 127–152.
- Van der Lans, R. M., Van de Grift, W. J. C. M., & Van Veen, K. (2018). Developing an instrument for teacher feedback: Using the Rasch model to explore teachers' development of effective teaching strategies and behaviors. *The Journal of Experimental Education, 86*(2), 247–264.
- Van Geel, M., Keuning, T., Frèrejean, J., Dolmans, D., Van Merriënboer, J., & Visscher, A. J. (2019). Capturing the complexity of differentiated instruction. *School Effectiveness and School Improvement, 30*(1), 51–67.
- Vangrieken, K., Meredith, C., Packer, T., & Kyndt, E. (2017). Teacher communities as a context for professional development: A systematic review. *Teaching and Teacher Education, 61*, 47–59.
- Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education, 24*, 80–91.
- Vogt, F., & Rogalla, M. (2009). Developing adaptive teaching competency through coaching. *Teaching and Teacher Education, 25*, 1051–1060.
- Webster-Wright, A. (2009). Reframing professional development through understanding authentic professional learning. *Review of Educational Research, 79*(2), 702–739.
- Wolthuis, F., Van Veen, K., De Vries, S., & Hubers, M. D. (2020). Between lethal and local adaptation: Lesson study as an organizational routine. *International Journal of Educational Research, 100*, 1–12.
- Xu, H., & Pedder, D. (2015). Lesson study: An international review of the research. In P. Dudley (Ed.), *Lesson study. Professional learning for our time* (pp. 29–58). Routledge.

Tijmen M. Schipper works as an associate professor Lifelong Learning and Development within the Business, Media and Law faculty of Windesheim University of Applied Sciences in Zwolle, the Netherlands. Tijmen was the first Dutch PhD student in the field of Lesson Study and obtained his PhD in 2019 on the dissertation 'Teacher professional learning through Lesson Study. An examination of Lesson Study in relation to adaptive teaching competence, teacher self-efficacy, and the school context'.

Sui Lin Goei was trained as a school psychologist and holds a PhD in instructional psychology. She is a well-known expert in inclusive education and how to cope with behavioral dilemma's within classroom teaching. She uses the collaborative professional development approach of Lesson Study to design lessons and interventions for inclusive and adaptive teaching. Together with dr. Siebrich de Vries and dr. Nellie Verhoef she has founded the Dutch Lesson Study consortium. Currently she divides her work between Windesheim University of Applied Sciences and VU Amsterdam, both in the Netherlands. At Windesheim she is professor of Inclusive Learning Environments and at VU she is an assistant professor in the academic group Learning Sciences and at the Teacher Academy.

Siebrich de Vries did her PhD in the field of teacher professional development. In 2013, she introduced Lesson Study in secondary schools and teacher training in the Northern part of the Netherlands. In 2016, she founded the consortium Lesson Study NL together with dr. Sui Lin Goei and dr. Nellie Verhoef. Siebrich currently works as a professor of applied sciences of Subject Pedagogy at NHL Stenden University of Applied Sciences, and as an assistant professor in the Teacher Education department at the University of Groningen where she supervises several PhD candidates in the field of Lesson Study.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 33

Adapting Teaching to Students' Needs: What Does It Require from Teachers?



Marieke van Geel, Trynke Keuning, Kyra Meutstege, Jitske de Vries,
Adrie Visscher, Christel Wolterinck, Kim Schildkamp, and Cindy Poortman

Abstract Teachers are increasingly expected to adapt their teaching to students' needs. This can be done by implementing differentiated instruction (DI) or assessment for learning (AfL). These concepts are regarded as two distinct approaches to identifying students' needs and adapting instruction accordingly. In the current study, we aim to identify empirical similarities and differences in teacher knowledge and skills required for differentiated instruction and assessment for learning respectively. Based on combined insights from two cognitive task analyses (CTA's), it appears that – in line with many other aspects of effective teaching – four phases are closely related for the task (either DI or AfL) as a whole: preparing a lesson series, preparing a lesson, enacting a lesson and, after this enactment, evaluating a lesson. The teacher skills required for DI and/or AfL in each of these phases are similar, however, the emphasis given to each skill differs in practice and this can be noted throughout all four interrelated phases. For AfL, the emphasis is on eliciting evidence during the lesson, for DI, the emphasis is on pro-active alignment of instruction and activities, based on students' needs. Since teachers need the same *underlying skills* to be able to perform either DI or AfL, we can hypothesize that teachers who are proficient at either DI or AfL, will also be able to develop and implement AfL or DI in practice.

Keywords Differentiated instruction · Assessment for learning · Teacher skills · Cognitive task analysis

M. van Geel (✉) · K. Meutstege · J. de Vries · A. Visscher · C. Wolterinck · K. Schildkamp · C. Poortman
Department of Teacher Development, University of Twente, Enschede, Netherlands
e-mail: marieke.vangeel@utwente.nl

T. Keuning
University of Applied Sciences KPZ, Hogeschool KPZ, Netherlands

1 Introduction

An important precondition for effective teaching is that teachers continuously try to obtain a valid picture of the extent to which their students are progressing towards the learning objective(s), and adapt their teaching based on that picture. Two common approaches to adapting teaching to students needs are differentiated instruction (DI) and assessment for learning (AfL). Differentiated instruction can take place by tailoring resources, methods of teaching, requirements for student outcomes, activities for learning, and curricula to suit the student's readiness, their learning interest or their learning preference (Smale-Jacobse et al., 2019; Tomlinson et al., 2003). DI "is a philosophy of teaching rooted in deep respect for students, acknowledgment of their differences, and the drive to help all students thrive" (Smale-Jacobse et al., 2019, p. 1). With DI, students will be challenged in areas they are strong in while receiving support in areas they are weaker in (Corno, 2008).

There are different approaches to DI and effects of these vary. However, in their meta-analysis Deunk et al. (2018) found that DI has an overall small positive effect on student achievement in primary education. A similar study revealed there are not many well-designed DI studies in secondary education, but the ones that were found showed small to medium effects of DI on student outcomes (Smale-Jacobse et al., 2019). The aforementioned 'different approaches' can take place both between and within classes.

The implementation of Assessment for Learning, defined as "encompassing all those activities undertaken by teachers, and/or students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged" (Black & Wiliam, 2010, p.7). These 'modifications' are "decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited." (Black & Wiliam, 2009, p. 9). If teachers apply AfL in the classroom, this can lead to higher student achievement (e.g., Kingston & Nash, 2015). The effectiveness of AfL is due to its high focus on continuous short feedback loops as both teacher and student are more aware of the current status of students in their learning progress, and of the next steps to take for students to achieve more learning objectives (Black & Wiliam, 2018).

In previous empirical research, we have investigated the knowledge and skills teachers need to implement DI and AfL separately. In the current study, we will combine insights from theory and practice, in order to identify similarities and differences between DI and AfL with respect to required teacher knowledge and skills, and factors related to the (perceived) complexity of providing DI and implementing AfL. These insights can be used to optimize coherence in the implementation of both approaches, separately or simultaneously, in order to enhance effective teaching by adapting education to students' needs.

2 Theoretical Framework

2.1 *Skills and Strategies for Differentiated Instruction*

Van Geel et al. (2019) identified and sorted skills and strategies required for the implementation of DI based on an analysis of instruments that are used to measure DI. The first three categories concern aspects that take place before the instruction, categories four and five during instruction and the last category is about more general teaching.

The first category is mastering the curriculum, which means that teachers need to have sufficient pedagogical content knowledge (PCK). PCK refers to subject-matter content knowledge, as well as knowledge about how to teach subject-matter knowledge. This means that teachers need to know how to teach students with differences in cognitive abilities and be aware of the effects of different classroom practices for weak, average, and high ability students (Deunk et al., 2015). Second is the identification of instructional needs through the analysis of assessments (van Geel et al., 2019). This can be done, for example, through pre-assessment in which teachers assess the degree to which students already master the learning objectives and to identify students' prior knowledge (Smale-Jacobse et al., 2019).

Next, the teacher needs to be able to, based on the identified instructional needs, set appropriately challenging learning objectives for all students. To do so, teachers need to have insight into performance goals on different levels (Deunk et al., 2015) and be knowledgeable about the domain they are teaching. The fourth category is monitoring: the teachers should monitor the students' progress and achievement (van Geel et al., 2019). Teachers do this by asking questions, observing students, checking students' work, using tests, etc. Monitoring should happen continuously and not at fixed moments in time (Smale-Jacobse et al., 2019) and teachers should use the insights to identify students' current level of learning and understanding (Deunk et al., 2015).

Fifth, teachers should adapt their instruction, materials, and assignments for students of different ability levels (Deunk et al., 2015; van Geel et al., 2019). This should be based on what they have monitored (van Geel et al., 2019), and as learning needs change (which will be discovered through the continuous monitoring in step four), the adaptations should be updated accordingly (Smale-Jacobse et al., 2019). Sixth, and finally, there are also general teaching dimensions such as realizing a safe and motivating learning environment or teaching students specific skills. Good classroom management and students feeling safe, welcomed, and respected are important preconditions for DI (Smale-Jacobse et al., 2019).

2.2 *Skills and Strategies for Assessment for Learning*

The implementation of AfL in the classroom requires the coherent and cyclical use of several strategies and skills (Veugen et al., 2021), aimed at identifying where the learner is going, where the learner is, and how to get from where the learner is to

where they should be going. Black and Wiliam (2010) identified five categories of AfL-skills (Black & Wiliam, 2010). Teachers should: (1) identify, clarify and share learning intentions; (2) engineer effective discussions, tasks and activities that elicit evidence of learning, (3) provide feedback that moves learners forward. Furthermore, students have an active role – teachers should (4) activate students as learning resources for their own learning as well as (5) for the learning of their peers.

When applying AfL, teachers determine what the learning objectives are for lesson(series) in order to establish what a teacher intends for students to learn in a lesson (Wiliam, 2011). In order to do this well, it is important that teachers have sufficient pedagogical content knowledge, which helps them to think about which learning objectives and corresponding learning tasks are appropriate for specific groups of learners. These learning objectives are complemented by success criteria: parameters that indicate where students are with regard to meeting the learning objectives. Teachers can clarify the learning objectives and criteria for success for example through dialogue with students (Carless & Boud, 2018). This can mean that teachers together with students look at and discuss examples of end-products previously completed by students (i.e.. ‘exemplars’).

After clarifying the learning objectives, teachers can elicit evidence on students’ learning progress and identify possible misconceptions through various assessment techniques, varying from more informal assessment techniques (e.g., on-the-fly observations or questions) to more formal assessment techniques (e.g., diagnostic tests). It is important to note that students can play an important role in eliciting evidence of their learning through self- or peer-assessment. Teachers may, for example, ask students to rate their own or each other’s work based on earlier established criteria for success.

Based on the evidence that the teacher elicited through assessment techniques, the teacher can stimulate student learning by giving feedback or adapt instruction based on the evidence. The effect of feedback, however, is very dependent on the context in which it is given (Shute, 2007). When AfL remains teacher-centered, students lack insight in learning objectives and are unable to interpret feedback in a meaningful way (Brooks et al., 2021; Nicol & MacFarlane-Dick, 2006). Next to giving feedback, teachers can also decide to redirect their teaching efforts (Kippers et al., 2018). Through eliciting evidence, teachers may have established misconceptions in students’ thinking regarding a certain topic or task. Instead of just asking students to re-try or re-think their solution, teachers may choose more fitting instructions, such as a worked example with a specific focus on the misconception.

Stimulating student agency in their own learning process is one of the key features of AfL. “Student agency refers to the quality of students’ self-reflective and intentional action and interaction with their environment.” (Klemenčič, 2015, p. 1). This can, for example, take the form of students formulating the criteria for success, or students that give each other peer feedback based on these criteria (Nicol & MacFarlane-Dick, 2006). Student agency is considered essential to the feedback literacy of students (Boud & Molloy, 2013). With increased student agency, students are more likely to be receptive to use feedback to redirect their learning efforts.

2.3 Combining Differentiation and Assessment for Learning

On the surface, DI and AfL may seem like quite different strategies: where AfL seems to emphasize the focus on gathering information (“assessment”) to use as feedback, in DI the adaptation of the instruction is emphasized. However, to make the assessment in AfL ‘for learning’ or ‘formative’, the teacher should actively do something with the information they gather, such as adapting the instruction (Wiliam, 2011). Likewise, for a teacher to adapt their instruction to the learning needs of the students in DI, the teacher starts with determining what the learning needs of the students are by monitoring or gathering information (van Geel et al., 2019; Smale-Jacobse et al., 2019). The similarity in DI and AfL can most prominently be noticed in the importance of goal-orientation and evidence-informed decision-making. In both DI and AfL, teachers formulate explicit goals and deliberately design the teaching and learning activities with the aim of reaching these goals, taking differences between students into account. Assessing and monitoring students’ progress and understanding is essential to inform teachers’ decision-making with regard to the adaptation of these teaching and learning activities.

However, it remains yet unclear what applying DI or AfL in the classroom requires from teachers. The current study was therefore aimed at identifying the empirical similarities and differences between teacher skills and knowledge necessary for implementing DI and AfL, and identifying factors related to the (perceived) complexity. Although students and student ownership play an important role in both DI as well as AfL, since this chapter is focused on what adapting to students’ needs requires from teachers, the focus is on the teacher.

3 Method

3.1 Context of the Study

In this chapter, we compare and combine insights from two studies: one into knowledge and skills secondary school teachers need to implement differentiated instruction, one into knowledge and skills required for the implementation of assessment for learning. Both these studies took place in secondary education in the Netherlands, where students enter secondary school around the age of 12 years. The Netherlands is known for a tracked system, students are assigned to a specific track based on their primary school performance. Three different tracks exist: pre-vocational (4-year program), senior general (5-year program), and pre-university (6-year program) (EP-Nuffic, 2015). In general, Dutch schools have a lot of autonomy, almost all decisions with regard to teaching, learning, and curriculum are made at the school level (OECD, 2008, 2010). Only at the end of their secondary education, students take part in national assessments (OECD, 2008). In general, secondary school teachers have a lot of freedom to shape their instruction.

3.2 *Cognitive Task Analysis Procedure*

Both DI as well as AfL are all about adapting teaching to students' needs. In the current study, we aim to identify what adapting teaching to students' needs requires from teachers. From previous research (e.g. van Geel et al., 2019) we know that providing differentiated instruction requires knowledge and skills that cannot be directly observed. In order to identify, analyze, and structure the skills and knowledge used by experts during the performance of a complex task a cognitive task analysis (CTA) can be performed (Clark, 2014). In this chapter, we therefore combine the outcomes of two CTA's that were performed to identify knowledge and skills required, one for the complex task of implementing AfL and one for the complex task of providing DI. In both CTA's, the steps as described by Clark et al. (2008) and refined by Van Geel et al. (2019) were applied: (1) collect preliminary knowledge, (2) identify knowledge representations, (3) apply focused knowledge elicitation methods, (4) analyze and verify data acquired, (5) format the results for the intended application.

In line with Van Geel et al. (2019), it was decided that the representation (step 2) would be (a) an overview in which all constituent skills, including the relationships between those skills are presented (also called: skill hierarchy) (b) an overview of the required knowledge to perform these skills, and (c) factors related to complexity of performing the task. In the two CTA studies, collection and analysis of data took place in an iterative process, where each stage of data collection was followed by a brief analysis, providing input for the next stage. In both CTA's, classroom observations were followed by semi-structured stimulated recall interviews. The CTA researcher asked the teacher to elaborate, in order to gather as much information as possible. In each CTA, after all interviews were conducted, an expert meeting was organized with the expert teachers as participants. In these expert meetings, a preliminary version of the skill hierarchy for the skill under investigation was developed and discussed. Next, content experts were consulted to verify and expand the findings from the previous steps. Both CTA's resulted in a skill hierarchy, including a detailed description of each skill and the desired level of performance (also called 'performance objectives'), and an overview of required knowledge. The CTA outcomes will be compared in order to identify similarities and differences between DI and AfL in practice.

3.3 *CTA Participants*

3.3.1 **Participants CTA Differentiated Instruction**

The focus in the CTA for DI was on mathematics. Eleven teachers, together teaching all levels and age groups of secondary education, participated in the classroom observations and stimulated recall interviews. Six of those teachers also participated

in the teacher expert meeting. Ten content experts (teacher educators, educational consultants, researchers and educational inspectors) participated in the second expert meeting.

3.3.2 Participants CTA Assessment for Learning

The CTA for Assessment for Learning was aimed at three secondary school subjects: English, Dutch, and chemistry. This focus was decided upon because these two languages are core curriculum, and chemistry is an important STEM subject (as well as the area of expertise of one of the researchers). Eight teachers (four for Dutch, two for English, two for chemistry) were each observed and interviewed for two lessons. Twelve teachers, of which four were also observed and interviewed, participated in the expert teacher meeting. In the content expert meeting, eight consultants and researchers participated.

3.4 Data Analysis

For the purpose of this chapter, a team of researchers (the first four authors of this chapter) discussed the findings from the two CTA's in order to identify similarities and differences between the skills required for DI and AfL. In this analysis, the labels, descriptions and performance objectives for each constituent skill were compared. The research team also compared the required knowledge and identified complexity factors for DI and AfL.

4 Key Findings

4.1 Skills

Although the wording in the two initial skill hierarchies differed, in-depth discussions and desired performance as described in performance objectives revealed striking similarities between the outcomes of the two separate CTA's. In Fig. 33.1 the two skill hierarchies of DI-instruction and AfL are therefore combined. In a skill hierarchy, constituent skills at lower levels enable the learning and performing of skills higher up in the hierarchy (e.g., Van Merriënboer & Tjiam, 2013). So, for example: in order to prepare a lesson series, it is required to be able to make a planning of a lesson series, and for planning a lesson series, it is required to be able to determine objectives. As can be seen in this overarching skill hierarchy, four phases that are closely related play an essential role for the task (DI or AfL) as a whole: preparing a lesson series, preparing a lesson, enacting a lesson and, after this enactment,

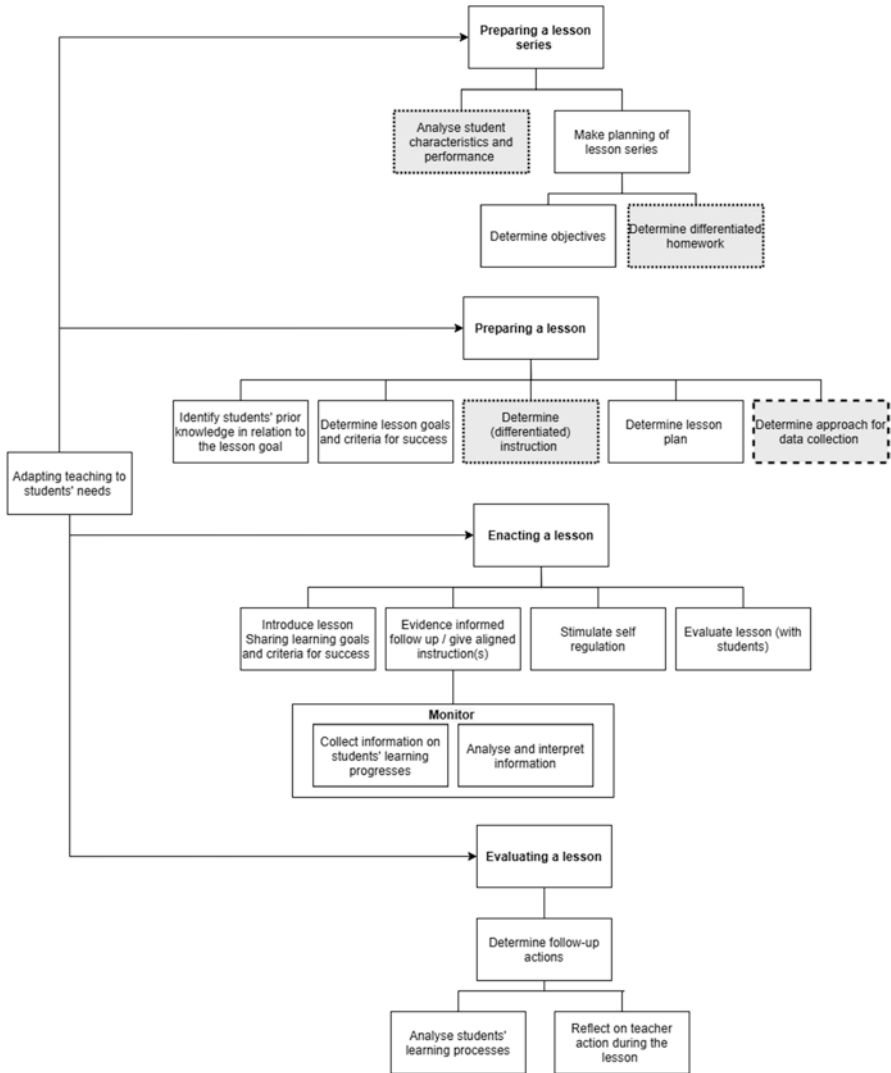


Fig. 33.1 Combined skill hierarchy for adapting teaching to students' needs
Note that skills represented with dotted lines exclusively stem from the CTA into DI, and the skill represented with dashed lines exclusively stems from the CTA into AfL.

evaluating a lesson. For teachers to be able to apply either AfL or DI, these four phases cannot be separated and seen as isolated activities. Coherence between the four phases is necessary for high-quality performance of the task as a whole.

Although the majority of skills appears similar across both AfL and DI, several skills are DI-specific (represented with dotted lines in Fig. 33.1) or AfL-specific (represented with dashed lines in Fig. 33.1). For both AfL as well as DI, teachers

need to prepare a lesson series. In order to do so, they make a planning (including *differentiated homework* for DI, e.g. teachers determine in advance which homework is suitable for challenging high-performing students and which homework will help low-performing to achieve the learning objectives) and determine objectives. For DI, the *analysis of student characteristics and performance* is also required in this preparation phase. This skill was not identified in the CTA for AfL. An explanation could be that for DI, teachers obtain a picture of their students' needs and progress for long-term preparation and possible adjustments in objectives. In the *lesson preparation phase*, both for AfL as for DI, teachers identify students' prior knowledge related to the lesson goal.

In the *lesson preparation phase*, one DI-specific and one AfL-specific skill were identified. For DI, teachers prepare differentiation instruction, they for example determine specific approaches to explaining the subject matter for high, average and low performing students. For AfL on the other hand, teachers specifically determine approaches for data collection: how will they, during the lesson, elicit information about students' progress, understanding, and/or misconceptions? This is strongly connected to the 'monitoring' skill during the lesson. However, teachers in the CTA for DI, did not explicitly mention that they prepare *how* they will monitor student understanding and progress during the lesson, whereas this was an explicit part of lesson preparation for teachers in the CTA study into AfL.

As can be noted from Fig. 33.1, during the phases *enacting a lesson* and *evaluating a lesson*, no AfL- or DI-specific skills were identified. This does not imply that AfL and DI are exactly the same, however, it does indicate that teachers need the same *underlying skills* to be able to perform either AfL or DI. A subsequent conclusion could be that teachers who are proficient at either DI or AfL, would probably also be able to perform the other task. Although the underlying required skills are similar, the emphasis given to each skill differs in practice and this can be noted throughout all four interrelated phases. For AfL, the emphasis is on eliciting evidence during the lesson. Teachers prepare their approach to data collection, during the lesson they analyze and interpret the information in order to utilize the insights for evidence-informed follow-up. For DI, the emphasis is on pro-active alignment of instruction and activities, based on students' needs. In order to do so, teachers collect information about their students' progress and understanding both in the preparation of a lesson series, and the preparation of a lesson, as well as by monitoring during the lesson. In general, it appears that students have a more active role in classrooms where teachers apply AfL. Although stimulating students' self-regulation in DI is also an important skill, the emphasis in DI is more on a pro-active approach by the teacher.

4.2 Required Knowledge

In both CTA's, next to required skills, required knowledge was identified. From the CTA into DI, three types of knowledge emerged: knowledge about students, subject matter knowledge, and general didactical-pedagogical knowledge. Basic elements

of teacher knowledge that were identified to be critical for applying AfL successfully are: domain knowledge, pedagogical content knowledge, knowledge of students' previous learning, and knowledge of assessment.

Knowledge about students (DI) is strongly related to knowledge of student's previous learning (AfL), although teachers in the CTA for DI stressed that it is not only of utmost importance to know about students' learning and performance, but also have insights into students' pedagogical needs. From the description of required subject matter knowledge (DI), it becomes clear that this encompasses domain knowledge (AfL) and pedagogical content knowledge (AfL). This knowledge is needed for teachers to be able to respond adequately to e.g. students' misconceptions and identify students' next steps in their learning process (Heritage, 2010). From the CTA into AfL, it was concluded that teachers need specific knowledge about assessment, various techniques for eliciting information, and how to apply these. From the CTA into DI, it appeared that teachers need general pedagogical didactical knowledge.

4.3 Factors Related to Complexity

It is generally assumed that adapting teaching to students' needs is a complex teaching skill. In order to support teachers in developing skills for adapting their teaching to the needs of their students, it is recommended to identify, and if possible: adapt, the external factors that influence the perceived complexity. This way, a sort of scaffolding is applied (Van Merriënboer & Kirschner, 2018): by providing teachers the opportunity to start with implementing DI or AfL in a less complex situation, they can focus on developing the skills necessary for DI or AfL. When teacher are able to apply their skills in a relatively less complex situation, the complexity of the situation can be increased. Since this (perceived) level of complexity of differentiated instruction and assessment for learning differs across situations (Van Geel et al., 2019), in the two studies expert teachers were asked to identify these factors related to complexity. In both studies, the same four general factors related to complexity were identified:

1. Student group composition: number of students, degree of diversity, classroom climate, students' task-orientedness.
2. Lesson content: topic, goal.
3. Curriculum material: assignments at different levels, diagnostic value of suggested instructional material, suggestions for remediation.
4. School support: facilities, duration of classroom hours, collaboration, testing structure and rules.

This list of complexity factors can provide a basis for developing a scaffolded professionalization trajectory, in which (beginning) teachers are encouraged to start implementing DI or AfL in situations with relatively low complexity, e.g. when teaching a rather easy topic to a rather homogeneous group of students.

5 Conclusion and Discussion

In this chapter, we aimed to identify empirical similarities and differences in required teacher knowledge and skills for adapting teaching to students' needs, by applying either assessment for learning or differentiated instruction. Studies into DI and AfL so far, seem to be mostly conducted separately, using their own terminology. However, based on the comparison of underlying skills and knowledge, required for either DI or AfL, identified by means of cognitive task analyses, it appears that teachers roughly need the same *underlying skills and knowledge* to be able to perform either DI or AfL. We can therefore hypothesize that teachers who are proficient at either DI or AfL, will also be able to develop and implement AfL or DI in practice. Since also in practice, there is an overlap in applied skills and strategies, it could also be assumed that teachers who apply AfL, differentiate their instruction based on the identified differences, or that teachers who apply DI, use AfL strategies to identify their students' needs.

We argue that the fields of DI and AfL and differentiation would benefit from greater integration to be able to reach the common goal of improved learning and achievement. Both approaches not only require largely the same underlying skills, they also complement each other. Teachers who would like to adapt their teaching to students' needs could benefit from combining the knowledge and skills required for DI and AfL. For example, teachers who are proficient in providing DI could strengthen their monitoring by explicitly determining approaches to data collection in their lesson preparation. On the other hand, teachers who implement AfL could improve their preparations by also analyzing student characteristics, and preparing differentiated instructions in order to be better able to adapt their teaching on the spot.

Since adapting teaching to students' needs is an important characteristic of effective teaching, both pre-service as well as in-service teachers could benefit from professional development activities aimed at enhancing the coherent combination of DI and AfL. The identified knowledge and skills required for high-quality integration of DI and AfL, from preparation to evaluation, can serve as basis for developing such (continuous) professional development programs.

References

- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- Black, P., & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 92(1), 81–90.
- Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551–575.
- Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education*, 38(6), 698–712.
- Brooks, C., Burton, R., van der Kleij, F., Ablaza, C., Carroll, A., Hattie, J., & Neill, S. (2021). Teachers activating learners: The effects of a student-centred feedback approach on writ-

- ing achievement. *Teaching and Teacher Education*, 105, 103387. <https://doi.org/10.1016/j.tate.2021.103387>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315–1325.
- Clark, R. E. (2014). Cognitive task analysis for expert-based instruction in healthcare. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 541–551). Springer.
- Clark, R. E., Feldon, D. F., van Merriënboer, J. J., Yates, K. A., & Early, S. (2008). Cognitive task analysis. In J. M. Spector, M. D. Merrill, J. J. G. van Merriënboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd. ed., pp. 577–593). Lawrence Erlbaum Associates.
- Corno, L. Y. N. (2008). On teaching adaptively. *Educational Psychologist*, 43(3), 161–173. <https://doi.org/10.1080/00461520802178466>
- Deunk, M., Doolaard, S., Smale-Jacobse, A., & Bosker, R. J. (2015). *Differentiation within and across classrooms: A systematic review of studies into the cognitive effects of differentiation practices*. Rijksuniversiteit Groningen.
- Deunk, M. I., Smale-Jacobse, A. E., de Boer, H., Doolaard, S., & Bosker, R. J. (2018). Effective differentiation practices: A systematic review and meta-analysis of studies on the cognitive effects of differentiation practices in primary education. *Educational Research Review*, 24, 31–54. <https://doi.org/10.1016/j.edurev.2018.02.002>
- EP-Nuffic. (2015). *Education System the Netherlands; the Dutch Education system described*. Nuffic. <https://www.epnuffic.nl/en/publications/education-system-the-netherlands.pdf> Retrieved June 2019
- Heritage, M. (2010). *Formative assessment: Making it happen in the classroom*. Corwin Press.
- Kingston, N., & Nash, B. (2015). “Formative assessment: A meta-analysis and a call for research”: Erratum. *Educational Measurement: Issues and Practice*, 34(2), 55. <https://doi.org/10.1111/emip.12075>
- Kippers, W. B., Wolterinck, C. H., Schildkamp, K., Poortman, C. L., & Visscher, A. J. (2018). Teachers’ views on the use of assessment for learning and data-based decision making in classroom practice. *Teaching and Teacher Education*, 75, 199–213. <https://doi.org/10.1016/j.tate.2018.06.015>
- Klemenčič, M. (2015). Student involvement in university quality enhancement. In J. Huisman, H. de Boer, D. D. Dill, & M. Souto-Otero (Eds.), *The Palgrave international handbook of higher education policy and governance* Palgrave Macmillan. https://doi.org/10.1007/978-1-137-45617-5_28
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218.
- OECD. (2008). *Education at a glance 2008: OECD indicators*. Organisation for Economic Co-operation and Development (OECD).
- OECD. (2010). *Education at a glance 2010: OECD indicators*. Organisation for Economic Co-operation and Development (OECD).
- Shute, V. J. (2007). Focus on formative feedback. *ETS Research Report Series*, 2007(1), i-47.
- Smale-Jacobse, A. E., Meijer, A., Helms-Lorenz, M., & Maulana, R. (2019). Differentiated instruction in secondary education: A systematic review of research evidence. *Frontiers in Psychology*, 10, 2366. <https://doi.org/10.3389/fpsyg.2019.02366>
- Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., Conover, L. A., & Reynolds, T. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *Journal for the Education of the Gifted*, 27(2–3), 119–145. <https://doi.org/10.1177/016235320302700203>
- Van Geel, M., Keuning, T., Frèrejean, J., Dolmans, D., van Merriënboer, J., & Visscher, A. J. (2019). Capturing the complexity of differentiated instruction. *School Effectiveness and School Improvement*, 51–67. <https://doi.org/10.1080/09243453.2018.1539013>

- Van Merriënboer, J. J. G., & Kirschner, P. A. (2018). *Ten steps to complex learning: A systematic approach to four-component instructional design* (3rd ed.). Routledge.
- Van Merriënboer, J. J. G., & Tjiam, I. (2013). Development and teaching of complex skills in invasive procedures. In P. Lanzer (Ed.), *Catheter-based cardiovascular interventions: A knowledge-based approach* (pp. 173–186). Springer. <https://doi.org/10.1007/978-3-642-27676-7>
- Veugen, M. J., Gulikers, J. T. M., & Den Brok, P. (2021). We agree on what we see: Teacher and student perceptions of formative assessment practice. *Studies in Educational Evaluation, 70*, 101027. <https://doi.org/10.1016/j.stueduc.2021.101027>
- Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation, 37*(1), 3–14.

Dr. Marieke van Geel is an assistant professor at the Department of Teacher Development at the University of Twente in the Netherlands. Her research focuses on how teachers and school leaders can use a wide variety of data for instructional decision making, and on professional development in these areas. She is especially interested in knowledge and skills teachers need in order to be able to make sense and use of all data available to them, ranging from monitoring subtle signals students send during a lesson, to analyzing curriculum based tests, and interpreting outcomes on standardized assessment. Email: marieke.vangeel@utwente.nl

Dr. Trynke Keuning aims at bridging academic research with daily educational practice. She is a lecturer at the University of Applied Sciences KPZ where she teaches teachers and school leaders to conduct research in their own practice. Additionally she performs postdoctoral research into data-based decision making and differentiated instruction, but also on knowledge, skills and attitudes that various professionals with regard to education and care for children (aged 0–14) need for the implementation and continuation of interprofessional collaboration.

Kyra Meutstege studied cultural anthropology and educational science and now works a researcher at the University of Twente. In close cooperation with a regional secondary school board she is working on a project with regard to differentiated instruction for mathematics teachers. By means of a cognitive task analysis, the knowledge and skills a mathematics teacher in secondary education needs to give differentiated instruction and what factors influence the complexity of it were identified. Based on these insights, Kyra is now developing a professional development trajectory that will be implemented and evaluated in 2021–2023.

Dr. Jitske de Vries graduated cum laude for her bachelor in Psychology, followed by her master in educational science. She conducted her PhD at the University of Twente. In her dissertation, she evaluated two projects related to teacher professional development, aimed implementing assessment for learning in secondary schools: the InformED project in the Netherlands, and the FORMAS project in which she collaborated with international project partners from Cyprus, Greece, Belgium and The Netherlands.

Prof. Dr. Adrie Visscher is a full professor at the University of Twente and head of the Department of Teacher Development. In his research he investigates how the provision of various types of feedback to students, teachers and schools (e.g., classroom observation results, students' perceptions of teaching quality, students' ability growth, feedback to students who are working on assignments) can support the optimization of the quality of classroom teaching and student learning. As the receipt of such feedback often is the starting point for improvement-oriented actions his research also focuses on the characteristics of effective teacher professionalization.

Dr. Christel Wolterinck was a PhD student at the Department of Teacher Development at the University of Twente. Her research interests center on assessment for learning and enhancing teachers' professional development in the area of assessment and data use. She is also a school

leader, working at Marianum in Groenlo, a school for secondary education belonging to the Foundation Carmelcollege. Previously, she worked as a chemistry teacher for 14 years. Currently, her work focuses on implementing assessment for learning and data-based decision making, specific the continuing professional development of teachers in secondary education and stimulating practitioner research in the schools.

Prof. Dr. Kim Schildkamp is a full professor in the Faculty of Behavioural, Management, and Social Sciences of the University of Twente. Kim's research focuses on data-based decision making and formative assessment. She is a Fulbright scholarship recipient, which she used to study data use in primary and secondary education in Louisiana. She is the previous president of ICSEI (International Congress on School Effectiveness and Improvement). She has published widely on the use of (assessment) data and was, for example, editor of the book "Data-based decision making in education: Challenges and opportunities" and "The data team procedure: A systematic approach to school improvement.

Dr. Cindy Poortman is an associate professor at the University of Twente, the Netherlands. Her research and teaching focus on teacher and school leader professional development in Professional Learning Networks (PLNs). Examples are data teams, teacher design teams, and research practice partnerships. Leadership, sustainability, data use and assessment for learning are her themes of focus. Publications include data use papers and the book *The Data Team Procedure* (2018, Springer). She also co-edited *Networks for Learning* (2018, Routledge) and co-edits the *Professional Learning Networks Book Series* (2019-, Emerald). Cindy is one of the InformED project researchers.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 34

The Teacher's Turn: Teachers' Perceptions of Observed Patterns of Classroom Interaction



Nienke Smit, Marijn van Dijk, Kees de Bot, and Wander Lowie

Abstract Insight in the way verbal teacher-student classroom interaction unfolds during the language lesson is of crucial importance for effective teaching. Although classroom observational research is indispensable, it is unable to uncover underlying intentions or motivations for the observed behavior. Teacher cognition research seeks to address the relation between teaching practice and what teachers think. This study reports on the perceptions of a group of English as a foreign language teachers ($n = 57$) who were asked to reflect on results from a classroom observation study about EFL teacher-student interaction in a similar teaching context. A large majority (82%) of the respondents recognized the observed pattern of closed teacher questions and limited student responses. This majority indicated that student participation in their own lessons is similar to the observed lessons or lower. Respondents attributed the pattern of high teacher activity and low student activity to emotional factors rather than to students' proficiency levels, lesson content, lesson activities or motivational aspects. According to 51% of the respondents, making students feel more competent by focusing on formative evaluation might improve classroom interaction, whereas 18% of the respondents suggested that interaction could be improved by using different teaching materials.

Keywords Interaction · Affective factors · Observation · Language teaching

1 Introduction

The main goal of foreign language teaching is to prepare learners to use the language in formal and informal settings of social interaction in order to co-construct meaning (Council of Europe, 2001; Larsen-Freeman & Anderson, 2011; Thornbury,

N. Smit (✉)
Utrecht University, Utrecht, The Netherlands
e-mail: n.smit@uu.nl

M. van Dijk · K. de Bot · W. Lowie
University of Groningen, Groningen, The Netherlands

2011). The foreign language lesson can be viewed as a social setting in which teacher and learners engage in interaction around a certain topic, for instance derived from a text. A meta-analysis (Murphy et al., 2009) revealed that active student engagement in classroom discussions about a text promotes co-construction of meaning. However, these authors also state that the way in which classroom discussions are organized matters greatly. An important prerequisite for effective discussions is that the teacher does not dominate the discussion, but that there is room for students to express thoughts, ideas and feelings during classroom interaction (Murphy et al., 2009). According to Murphy et al. (2009), it is not so much the *quantity* but the *quality* of classroom discussions that matters greatly in achieving co-construction of meaning.

Many researchers have acknowledged the importance of fostering co-construction of meaning in the language classroom (Gibbons, 2015; Walqui & Van Lier, 2010; Mercer & Dörnyei, 2020). However, classroom dynamics may be influenced by a host of factors, for instance student ability, number of students in the classroom, lesson topic and type of classroom activities (Dewaele, 2020; Mercer & Dörnyei, 2020; Dörnyei et al., 2015). These factors might impact the extent to which co-construction of meaning between teacher and students is achieved. A recent observational study (Smit et al., 2022) focused on what teachers and learners do to foster co-construction of meaning during interaction and revealed a gap between what is happening in classrooms and what research says about effective classroom interaction. The study provided systematic descriptions of teacher and learner question and answer behavior, and operationalized co-construction of meaning as active participation in question and answer sequences by everyone in the classroom most of the time. Asking questions is one of the basic tools in a teacher's pedagogical repertoire (Murphy et al., 2009). A teacher's open-ended question (i.e. no predetermined answer) can serve as an invitation for learners to contribute to co-construction of meaning. Smit et al. (2022) found highly active teachers and rather inactive students.

An important question with regard to educational research and teaching practice is to what extent they might inform each other. Research findings are not always understood, recognized or deemed relevant by practitioners. The general aim of this study was to bridge the theory-practice gap. The observational study of Smit et al. (2022) did not reveal underlying factors for the observed behavior. The first aim of the present study was to find out whether teachers who were not observed but work in the same teaching context in The Netherlands, think the observational evidence is representative of actual practice. The second aim was to investigate how teachers in The Netherlands would attribute the observations and what they thought might improve teacher-student interaction patterns in EFL lessons in the Netherlands.

2 Literature Review

2.1 *English as a Foreign Language Teaching in the Netherlands*

English is one of the three core curriculum subjects in the Dutch curriculum for secondary education. Communicative foreign language teaching forms the backbone of the national curriculum for English as a foreign language (EFL) (Fasoglio et al., 2015). The Dutch curriculum has been aligned with the Common European Framework of Reference (hence CEFR) and requires from 15–18-year-old students that they are able to enter discussions about a wide range of both familiar and unfamiliar topics at CEFR level B1+ / B2 (Fasoglio et al., 2015; Council of Europe, 2001). Understanding texts also plays a major role in the Dutch curriculum. By the end of secondary education, Dutch learners in the highest levels¹ take a national standardized reading exam at CEFR level C1 (Fasoglio et al., 2015). This exam determines 50% of the final grade for English. These curricular requirements illustrate why it is important for Dutch teenagers to be able to read English texts and discuss these texts during foreign language lessons at school.

2.2 *From Observations to Perceptions of Classroom Interaction: The Role of Lesson Content, Teaching Materials and Language Proficiency*

Factors that have been suggested as a major influence on how the language lesson unfolds are lesson content (i.e. what is talked about), teaching materials, and learners' language proficiency (Thornbury, 2011; Larsen-Freeman & Anderson, 2011). Regarding the content of the language lesson, the discussion is complicated. In a language lesson any topic could be approached from a language learning perspective, but according to Arnold (1999) it is crucial that the subject matter is appealing and relevant to the language learners. In order to foster learner engagement, proposals have been made to incorporate learner-oriented topics in the lessons (Maley, 2011). However, when this was operationalized as using lesson content derived from popular culture (e.g. film, music, celebrities) to focus on grammatical structures, this did not automatically lead to increased learner engagement (Piggott, 2019; Lightbown, 2015; Dönszelmann et al., 2020).

¹Dutch secondary education is ability streamed. Students from the age of 12 onwards enter one of the three levels of secondary education: pre-vocational, general secondary education and pre-university education. The current study focuses on students in the highest two levels.

Considering the second factor, teaching materials, recent studies have shown that the coursebook determines what happens in Dutch EFL lessons (Tammenga-Helmantel & Maijala, 2019) and that a heavy focus on restricted language practice does not help learners to interact in real-time (Van Batenburg et al., 2018). Additionally, these studies revealed large amounts of cognitively and sometimes also linguistically unchallenging discourse (Van Batenburg et al., 2018; Tammenga-Helmantel & Maijala, 2019). This suggests a possible gap between the way coursebooks prepare students for social interaction and the skills that are needed for actual social interaction inside and outside the classroom.

Thirdly, in order to interact with other people in another language, sufficient lexico-grammatical knowledge as well as a sufficient level of oral fluency are needed (Council of Europe, 2018). A study of oral fluency levels of Dutch teenagers by Fasoglio and Tuin (2017) confirmed that students in the two highest levels of Dutch secondary education attain the desired proficiency level, i.e. CEFR B1-B2 for speaking. Moreover, this study showed that a large proportion (48.6%) of the students in pre-university education achieve CEFR C1 level for oral fluency. An important additional finding from this study was that Dutch teenagers, although fluent enough, often do not use the English language in the classroom. In a sample of teenagers in pre-university education ($n = 385$), 20% of the students reported never to attempt to only use English as the language of communication during classroom interaction. These results suggest that active classroom participation is not a precondition for students to achieve relatively high fluency levels. Only 10% of the students in pre-university education always try and use English during the language lesson. In lower levels of secondary education, the percentage of students who speak English in class was even lower (Fasoglio & Tuin, 2017). Although Dutch teenagers seem to be reasonably fluent in English, they show limited evidence using the language in the classroom.

2.3 From Observations to Perceptions of Classroom Interaction: The Role of Emotions and Motivation in Classroom Interaction

The fourth and fifth factor that might impact classroom interaction relate to affective aspects in the language learning process (Arnold, 1999). We will discuss both emotions and motivation in relation to Self-Determination Theory, hence SDT (Deci & Ryan, 2000). SDT focuses on what moves people into action by describing human psychological needs in terms of relatedness, autonomy and competence. Gibbons (2015) illustrates competence and relatedness by discussing the role of emotions and stresses that a certain amount of struggle in understanding others and making yourself understood is needed to get ahead in language learning. She also points out that moments of frustration are most significant when learners are communicating with “a helpful interactant” (Gibbons, 2015). However, when frustration causes

students to lose confidence and feel embarrassed or anxious, learning stalls. According to Dewaele et al. (2018) lessons which are emotionally uninteresting or emotion-free, might lead to routine, boredom and lack of engagement, which could suggest a weak sense of relatedness.

A student who is bored might try to avoid active participation, but a lack of response from the learners could in turn influence the teacher's sense of relatedness and competence, which in turn could affect interaction. Although proficiency levels of qualified English teachers in the Netherlands are at CEFR C1/C2 (10 Voor de Leraar, 2018) and there is no evidence that teacher proficiency might be a limiting factor, Dönszelmann (2019) reports that foreign language teachers confessed to struggle being consistent in their use of the foreign language during the lessons. Whereas linguistic competence might not be at stake, a threat to relatedness or experienced autonomy and teaching competence might play a role here. Underlying emotional factors for this struggle to use the English language consistently might be that teachers' worry that students do not understand what they are saying, or that students and parents might complain about the intelligibility of the language lesson (Fasoglio & Tuin, 2017; Dönszelmann, 2019).

Finally, learner motivation might also impact classroom interaction. Language learning motivation might fluctuate during the lesson and these fluctuations could impact the quality and quantity of student participation during the language lesson (Waninge et al., 2014). Research into language learning motivation has focused on factors such as the value and relevance for the language user, being able to use the language, and the goals learners want to achieve (e.g. educational or professional advantages) (Dörnyei et al., 2015). These factors also relate to SDT's relatedness and competence (Deci & Ryan, 2000), constructs which are closely associated with cognitive, emotional and behavioral engagement (Mercer & Dörnyei, 2020). National surveys revealed that Dutch teenagers have a positive attitude towards the English language and its relevance (Fasoglio & Tuin, 2017). Based on Dutch teenagers' self-reported levels of emotional engagement with the English language would suggest sufficient motivation to learn this language. However the multi-dimensional and dynamic nature of this construct (Waninge et al., 2014) might also implicate that sufficient motivation might not directly lead to active verbal student behavior during classroom interaction.

2.4 Observed EFL Classroom Interaction and Teacher Cognition

Teacher cognition research seeks to address the relationship between what teachers do in their teaching practice and what they think, know and believe. This type of research is often carried out to complement classroom observational research (Borg, 2006; Basturkmen, 2012). Johnson (2006) stresses that teacher cognitions and pedagogical decisions mutually influence each other and change over time. It is therefore

important to examine both teaching behavior, which is defined here as what teachers do during their lessons, and teachers' perceptions of the observed behavior.

Questions and answers are building blocks of social interaction that can be observed and labelled relatively clearly and were therefore chosen by Smit et al. (2022) as a representation of moment-to-moment teacher-student interaction patterns that occur naturally in a language lesson. The results from this observational study revealed that teacher questions and student answers have the tendency to form patterns dominated by closed teacher questions and simple student answers. During a 50-minute lesson, English as a foreign language teachers asked around 60 questions on average to which students gave short (i.e. one to three-word utterances) or no answers. Micro-level observations also revealed that in 30% of the lessons ($n = 16$), students had the tendency to adjust the level of their answer to the level of the teacher question (e.g. 'low level' questions leading to 'low level' answers, higher level questions leading to higher level answers). However, this study found no evidence for a relation between the teachers' follow-up question and the previous student answer. The study provided detailed descriptions of the micro-dynamics of teacher-student interaction in foreign language lessons, but did not yield insight in underlying reasons for the observed interaction patterns (Smit et al., 2022).

3 The Present Study

The first aim of the present study was to find out whether teachers think the observational evidence found in Smit et al. (2022) is representative of actual teaching practice. The second aim was to investigate how teachers would attribute the observed patterns and what they would suggest as directions to improve teacher-student interaction patterns in EFL lessons in the Netherlands. The present study was designed to minimize attribution errors that might be caused by the actor-observer effect of confirmation bias. Teachers may have varying reasons for choosing to participate in an observational research study. However, the presence of a camera in the classroom might influence teacher and student behavior, making it difficult to determine to what extent the observations are "business as usual". Therefore teachers from the same teaching context who had not been observed were asked to participate in this study. The study seeks to answer the following research questions:

1. Do teachers recognize the observed interaction patterns that are characterized by a dominance of closed teacher questions and short student answers in their own teaching practice?
2. What is the best explanation for the observed classroom interaction according to EFL teachers?
3. What do teachers perceive to be the best suggestion for improving teacher-student classroom interaction?

4 Method

4.1 *Participants and Context*

Teachers ($n = 57$) attending a presentation about classroom interaction were asked to participate in a short questionnaire about the classroom observational evidence. The data was presented and explained by the first author of this paper on two different occasions in January and March 2020. The first group of respondents ($n = 47$) were EFL teachers participating in a teacher conference organized by the University of Groningen in January 2020. One of the conference participants was not a teacher, but worked as a consultant for an educational publisher. This respondent was excluded from the study. The second group of respondents ($n = 10$) were trainee teachers in the Master of Education at the University of Groningen attending a seminar about interaction in the language classroom. This seminar was part of an English language teaching methodology course taught by the first author of this paper. During their masters' program the trainee teachers also worked as EFL teachers in schools for secondary education in the Netherlands.

All respondents in this study ($n = 57$) were familiar with the EFL teaching context in Dutch secondary education and had hands-on teaching experience. Respondents were asked to answer our questions as if it were their own practice. The response rate for completing the anonymous questionnaire was 100%, which might be due to the convenience sampling procedure described above and the short amount of time needed to complete the questionnaire (less than 3 min on average). All participants were first asked for consent to participate and were given the possibility to opt out immediately. The research design was approved by the Ethics Committee of the Department of Teacher Education at the University of Groningen (EC reference 19-024/RM/AA).

The sample consisted of respondents working in different levels of Dutch education. A large majority (86%) of the respondents was female, 12% were male and one person (2%) indicated "other" for gender. An overall majority of the respondents were EFL teachers working with teenagers in two highest levels of Dutch secondary education² (43 people – 74%), 13 teachers (24%) taught English in (pre)vocational secondary education (teenage learners), one person (2%) worked as an EFL teacher in higher education (young adult learners, >17 years old). The distribution between experienced and early career professionals (defined as anyone who had between 0 and 5 years of experience) was roughly two-thirds (35 people – 61%) to one third (22 people – 39%). This means that the majority of the respondents who reflected on the classroom observational evidence that was presented during the presentation had substantial experience teaching learners of a similar age and educational level (i.e. higher secondary and pre-university education).

²See footnote 1 for a brief explanation of Dutch secondary education.

4.2 Procedure

At the start of the presentation, the first author of this paper explained the relevance of classroom interaction research and provided some background information about the context of the research project. The teachers were informed that observational data in Dutch secondary education classrooms had been collected in lessons taught to learners (14–17 years old) preparing for higher vocational or university education. All observed lessons used a text as a language input, which meant that lessons with a focus on teaching grammar were excluded from this study. It was explained that classroom interaction had been studied by observing sequences of teacher questions and students' answers and that teacher questions and student answers had been coded with the Questions and Answers in English Language Teaching (QAELT) coding scheme (Smit et al., 2022). This coding scheme consists of four-point scales for teacher questions and student answers in which openness and level of complexity are accounted for. Table 34.1 displays the simplified version of QAELT coding scheme as presented to the respondents.

After explaining the coding system, the observational evidence was presented. For the representation of the observational data three State Space Grid visualizations (Hollenstein, 2013; Lamey et al., 2004) were used. The scale for teacher questions is displayed on the horizontal axis of the State Space Grid and the vertical axis displays the scale for student answers. Together these scales form a 4x4 grid. Every dot in the grid represents an interaction which is formed by a teacher question combined with a student answer. The respondents were first informed that the “closed question – simple answer” pattern was the dominant pattern for the majority of the observed lessons (5 out of 16 lessons, i.e. 31%). The closed question-simple answer cell is the region in which most interactions took place. Then a State Space Grid showing a lesson with high levels of interaction and a different type of dominant pattern was presented to the respondents. This was the state space grid of lesson d4 displayed in Fig. 34.1. The grid of lesson d4 reveals that the teacher received an answer to every question. Additionally, the majority of the questions in this lesson took place at the level of clarification or open-ended questions.

Next, the teachers looked at a lesson (a1) with a low level of interaction (see Fig. 34.2). In this lesson the closed question and the simple answer, indicated by the yellow box, was the dominant pattern. Notably, a lot of questions that were asked during this lesson did not receive an answer at all.

Finally, teachers gauged State Space Grid b2 (Fig. 34.3) which depicted the median level of observed interaction in EFL lessons from the data set that was used

Table 34.1 Simplified version of QAELT coding scheme

Teacher question	Code	Student answer
Non-elicitation	0	No response
Closed question	1	Simple
Clarification	2	Complete
Open-ended	3	Complex

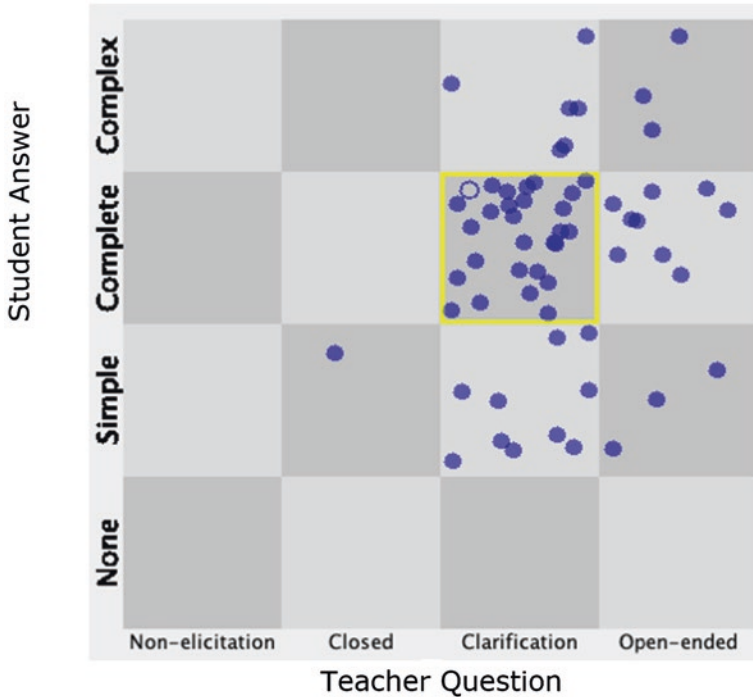


Fig. 34.1 State Space Grid visualization of a lesson with high levels of interaction (lesson d4)

in Smit et al. (2022). In order to establish the median level of teacher-student interaction the following measures were used: number of questions and percentage of questions in most the frequently occurring cell of the State Space Grid. The median number of teacher questions uttered during a 50-minute lesson in the dataset was 51. The most frequently occurring cell in this data set was the closed question – simple answer cell. The lesson with the median percentage of interactions (26%) in this cell was lesson b2. From a sample of 16 lessons, seven lessons had a lower percentage of interactions in the dominant cell and eight lessons had a higher percentage in the dominant cell. It was explained to the respondents that we chose to show the median level of observed interaction in order to validate the sample median. We asked the respondents whether they thought the level of interaction in their lessons was either lower or the same, or higher than the median level of interaction in the sample. It was explained to the respondents that lesson b2 represented a lesson “in the middle”, represented by the median.

The respondents filled out the digital anonymous Qualtrics (<https://www.qualtrics.com>) questionnaire immediately after the presentation. The questionnaire could be accessed by the participants by using a QR code or a shortened url. After filling out consent, gender, teaching experience and type of school in which the teachers worked, they were asked to answer the questions in Table 34.2 based on their expertise.

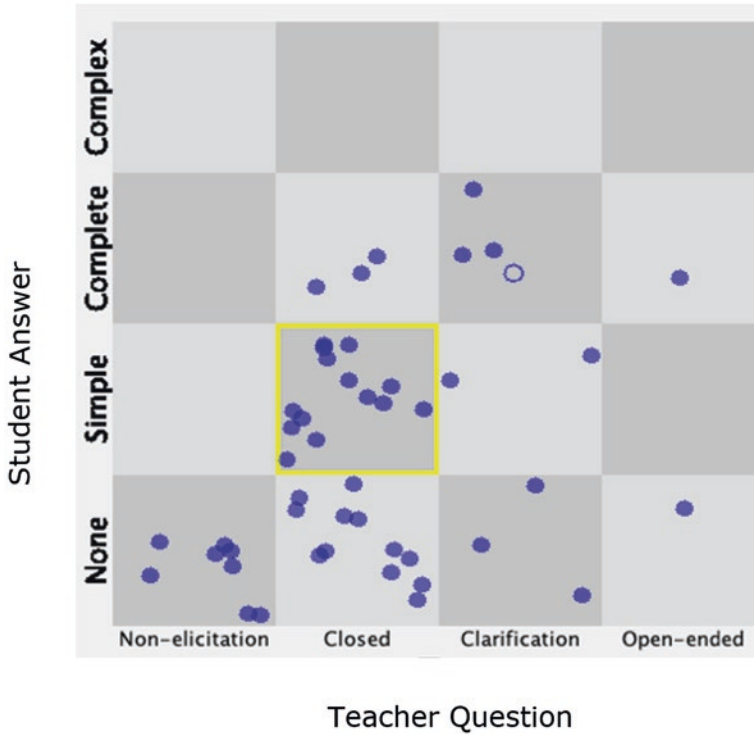


Fig. 34.2 State Space Grid visualization of a lesson with low levels of interaction (lesson a1)

The questionnaire was designed in such a way that there was a relation between the answer options to question 3 and the answer options to question 5. Question 3 consisted of possible explanations for the observation classroom observation patterns and question 5 consisted of possible measures for improvement aligned with the explanations. Table 34.3 shows how the answer options of these two questions correspond.

From Table 34.3 it can be seen that dedicating classroom attention to vocabulary and conversation skills was suggested in order to address possible language learning issues. Making students feel more competent (for instance by using formative evaluation techniques) was proposed to overcome possible emotional barriers. Problems in lesson content might be addressed by teaching about topics that students are interested in. A solution for teaching materials that do not encourage learners enough to participate actively would be to make teaching materials more interesting. And finally, motivational factors, for instance students who do not want to learn English at school, could be targeted by actively increasing students' motivation to learn English. Both questions 3 and 5 had a forced response, which means that participants were asked to pick only one explanation and only one measure.

Immediately after filling in the questionnaire, group results for all questions were displayed to the respondents, after which the first author of this paper and the respondents engaged in a brief discussion about the results. The goal of this discussion was teacher development and therefore not included in this study.

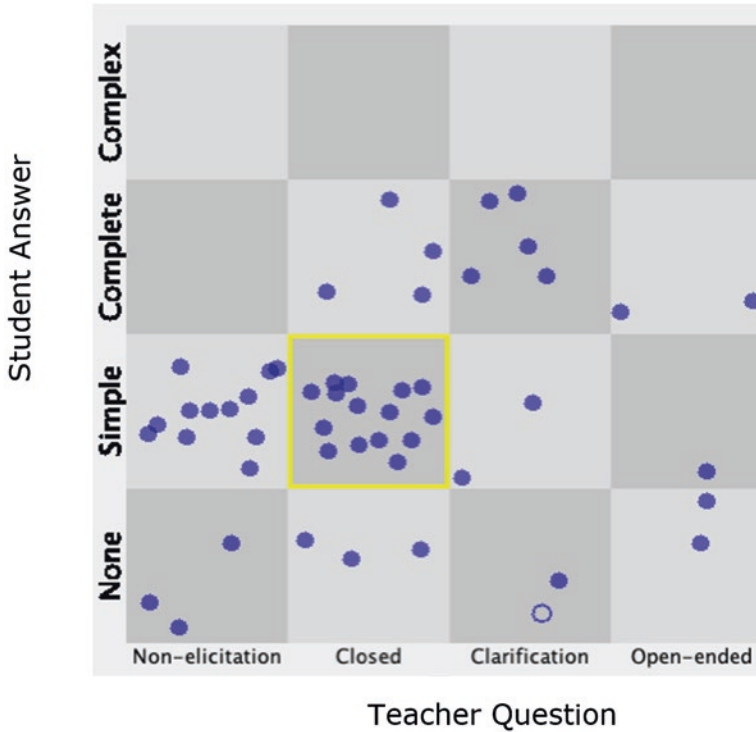


Fig. 34.3 State Space Grid visualisation of a lesson with the median level of interaction (lesson b2)

Table 34.2 Questionnaire about teachers’ perceptions

#	Question
1.	Do these observations confirm what you expected?
2.	How do you perceive the level of teacher-student interaction in your lessons?
3.	In your opinion, what is the best explanation for this type of classroom behavior?
4.	Do you think there is way in which this type of interaction can be improved?
5.	In your opinion, what is the BEST measure to improve classroom behavior. Please choose one option.
6.	OPTIONAL: Please write down any other ideas you have to encourage students to be more active during the language lesson (open question)

5 Results

Regarding the research question (RQ1) whether teachers recognize the dominant patterns of classroom interaction, an overall majority of the respondents (82%) confirmed that the observations were in line with their expectations. A small minority (7%) indicated that the results were worse than they had expected, and 11% indicated that this was better than they had expected. When the teachers were asked if they thought classroom could be improved, almost all respondents (96%) said ‘yes’

Table 34.3 Explanations for classroom interaction and possible measures to improve

Answer options question 5		#	Answer options question 6
1	NA	1	I do not think classroom interaction can be improved
2	Language skills: The students are not fluent enough	2	By increasing attention for vocabulary and conversation skills
3	Emotional factors: Speaking the foreign language in the lesson makes students feel uncomfortable	3	By making students feel more competent (formative evaluation)
4	Lesson content: Students are not interested in the lesson topic,	4	By teaching about topics that interest the students
5	Teaching materials: Teaching materials do not encourage students to participate actively,	5	By making teaching materials more interesting
6	Motivational factors: Students do not want to learn English at school.	6	By increasing the motivation for learning English

and only two (4%) believed that improvement was not possible. Regarding the teachers' self-assessment of interaction patterns in their own lessons the results show that 72% of the respondents thought that the level of classroom interaction in their lessons is similar or lower to the observed median level of interaction. The results show that roughly a third (30%) of the respondents indicated that the level of interaction in their lessons is higher.

With regard to the question of what the best explanation for the most frequently observed patterns of classroom interaction was (RQ2), a majority (72%) attributed the observed interactions patterns to emotional factors (see Fig. 34.4). According to 14% of the respondents, a lack of encouraging teaching materials is the best explanation for the observed results. This means that most respondents suggested that emotional factors play an important role in the emergence of classroom interaction patterns that are characterized by active teachers asking many closed questions and inactive students giving no answers or very short answers.

Further analyses of the responses revealed that a large majority (81%) of the experienced (>5 years) teachers attributed the observed interaction patterns to emotional factors. A smaller majority (59%) of the inexperienced teachers (0–5 years) thought that emotional factors were the best explanation for the observed patterns. One in three (31%) inexperienced teachers mentioned that the content and teaching materials could be a possible explanation for relatively inactive learners.

Regarding the possibility for improvement (RQ3), 98% thought improvement was possible. The results of the follow-up question (Table 34.2, question 5) about measures to improve classroom interaction are displayed in Fig. 34.5. The proposed measures to improve classroom interaction were increasing attention for vocabulary and conversation skills, making students feel more competent (formative evaluation), teaching about topics that interest the students, making teaching materials more interesting and increasing the motivation for learning English. Making students feel more competent by using formative evaluation was the most promising measure according to the respondents (51%). Making teaching materials more interesting was also suggested (18%), one respondent (2%) thought that classroom

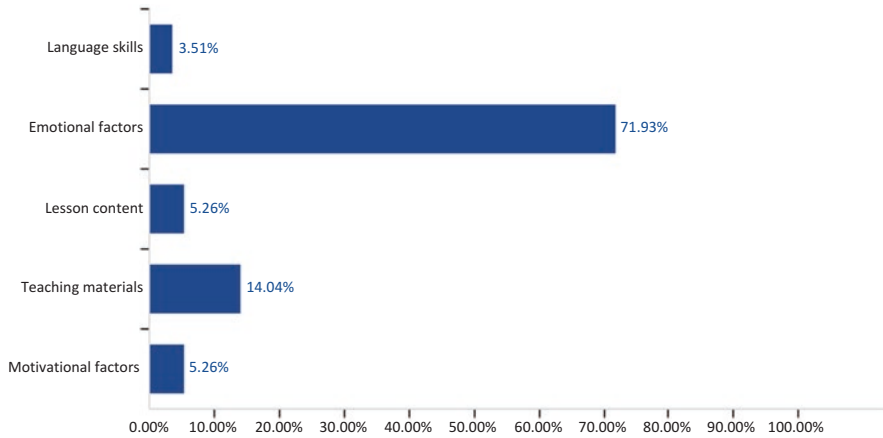


Fig. 34.4 Best explanation for classroom interaction according to the participants

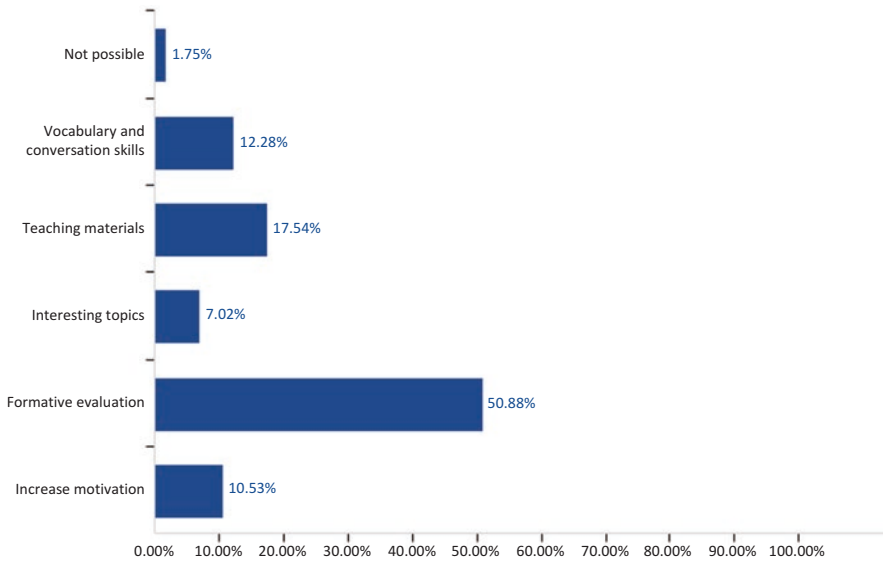


Fig. 34.5 Measures to improve classroom interaction

interaction could not be improved, incorporating more interesting topics were suggested by four teachers (7%), six teachers proposed increasing motivation (10%) and seven teachers (12%) preferred the option to improve vocabulary and conversation skills.

In the final question of the questionnaire teachers were also offered the opportunity to indicate how they thought classroom interaction could be improved. Nineteen respondents (33%) answered this question. The suggestions provided by the respondents could be linked to the following five broad themes: classroom organization

Table 34.4 Qualitative analysis of answers to the open question

Category	# comments	Examples
Classroom organization	9	"I think students are creatures of habit, and impeded by peer pressure. The best explanation for the questions and answers are, in my opinion, the product of habits. No matter how low a learner's proficiency, all of them are able to say "may I go to the toilet?". All of them can produce meaningful output. Creating new habits and expectations can solve this."
Curriculum	2	"Set up a collaboration with primary schools to encourage classroom interaction from an early age."
Lesson content	1	"Give students opportunities to choose their own topics."
Professional development	2	"I found it difficult to choose one answer because I think there are several reasons and it also depends on the students. I think I as a teacher could learn more about this."
Teaching materials	5	"There is a mix of factors that influence student interaction. I opted for the emotional aspect, but also see that this can be overcome by topic and material that interest students."

(47%), the national curriculum with a focus on starting early (11%), lesson content (5%), professional development (11%) and teaching materials (26%). Table 34.4 gives an overview of the themes, the number of comments made and for every theme one illustration of the answers given by the respondents.

Suggestions regarding improvements in classroom organization, especially the importance of a safe classroom climate were given most often as an additional solution for the lack of student activity. Teaching materials were mentioned by the respondents who opted for emotional factors in the closed question and who also indicated that more factors might play a role. Teaching materials were also mentioned in relation to using technology and digital tools.

6 Discussion and Conclusion

A group of EFL teachers who had not previously been observed were asked to reflect on observational findings on classroom interaction in their teaching context in The Netherlands. A very large majority of the respondents (82%) recognized the observed patterns, which could indicate that interaction patterns characterized by active teachers and inactive students might be a familiar struggle for many teachers in the Netherlands. The respondents were presented with observations of a lesson with a median level of interaction and we asked them whether their lessons had higher levels of interaction or the same or lower. Overall, respondents indicated that the observed interaction patterns confirmed their expectations of classroom dynamics regarding teacher questions and student answers. Only a third of the respondents thought that the level of active student participation during classroom interaction in

their lessons was higher, which could imply that average levels of active student participation in EFL lessons in the Netherlands might be somewhat lower than observed. A large majority of the respondents believed that classroom interaction can be improved.

From the literature, we know that joint attention and joint action are important mechanisms to achieve co-construction of meaning in the language classroom (Allwright, 1984; Larsen-Freeman & Cameron, 2008). Our respondents attributed the lack of students' responsiveness to teacher questions mainly to emotional factors. Some of the respondents suggested that a lack of student responsiveness might be due to classroom routines which are not conducive to language development. An example of such classroom routines are situations in which a language teacher asks questions which students can easily answer, or moments during the lesson in which teachers accept short answers. These asymmetric interaction patterns can be frustrating for teachers and potentially boring or uncomfortable for teenagers. According to Gibbons (2015) frequent interactions characterized by closed teacher questions and simple student answers could be characterized as a "high support/low challenge" interaction.

Whether teachers and learners actually are conscious of their own behavior (i.e. closed questions, simple answers) in real-time and the potential effect this might have on lessons, we do not know. A possible explanation might be that it is cognitively too demanding for teachers to monitor both a large group of students and themselves during the teaching-learning process. However, suggestions provided by the respondents indicate that teachers who might consciously or unconsciously work hard to maintain a safe learning climate, could also lead to routines in which teachers avoid putting teenagers on the spot by pushing for more extensive verbal output in English.

Learners who let their teacher to do most of the talking might implicitly shift the responsibility for managing the interaction to the teacher. From the perspective of teenage students, this might be an attractive option: limiting the amount of what you say can be an effective way to reduce risk of entering a potentially awkward, difficult or embarrassing situation in which you lose face in front of your peers. The benefits for teenagers of merely showing the teacher that they are "on board" by just listening and giving short but correct answers are high. This suggests that in whole class teacher-student interaction both learners and teacher could benefit from adhering to a relatively traditional distribution of authority. Future research, for instance observations of interpersonal behavior (Pennings et al., 2014) combined with a stimulated-recall interview, might look into whether the implicit agreement, the teacher leads and talks, whereas the students follow and answer, exists.

In order to overcome potentially uncomfortable situations, the respondents in this study offered some practical solutions such as asking questions but also using digital tools to let all students first give an anonymous online answer, before entering a classroom discussion. The respondents argued that this might lower the threshold for students. Adopting classroom management techniques to maximize active participation might offer suggestions to improve the balance between levels of teacher and student activity (Scrivener, 2012; Mercer & Dörnyei, 2020).

Research in the field of positive psychology suggests that fostering positive emotions can enhance language learning (MacIntyre et al., 2019; Dewaele, 2020). However, ignoring negative emotions like frustration, embarrassment and boredom and failing to address these might result in suboptimal behavioral patterns that are hard to change. Acknowledging that negative emotions are part of the learning process and offering opportunities to fail and learn from frustration might be needed to pave the way for positive experiences of learning and development and fostering relatedness (Gibbons, 2015; Deci & Ryan, 2000). In order to overcome suboptimal patterns of teacher-student interaction, a small majority of the respondents proposed to invest in formative evaluation practices. Formative evaluation is focused on getting ahead by providing ongoing interactive feedback during the *process* of learning. Process feedback might simultaneously address the basic needs of relatedness, autonomy and competence: helping students understand their current level of communicative competence, offering suggestions to change real-time behavior in order to become more autonomous, whilst helping each other in getting ahead by keeping the classroom conversation going.

It is promising that teachers recognize emotional struggles and suggest that researchers direct their attention to the cognitive and affective domain of learning simultaneously (The Douglas Fir Group, 2016). It is also promising that teachers express a wish to better understand and change classroom interaction. This study has shown that asking teachers to reflect on observational evidence of interaction patterns might improve their understanding classroom interaction and encourage them to reconsider how to make the most of the teacher's turn.

References

- 10 Voor de Leraar. (2018). *Kennisbasis Engels*. Vereniging Hogescholen. <https://kennisbases.10voordeleraar.nl/pdf/kennisbasis-vakmasters-engels.pdf> [Body of knowledge master of education in English, allied universities of applied sciences, The Netherlands].
- Allwright, R. L. (1984). The importance of interaction in classroom language learning. *Applied Linguistics*, 5(2), 156–171. <https://doi.org/10.1093/applin/5.2.156>
- Arnold, J. (Ed.). (1999). *Affect in language learning*. Cambridge University Press.
- Basturkmen, H. (2012). A review of research into the correspondence between language teachers' stated beliefs and practices. *System*, 40, 282–295. <https://doi.org/10.1016/j.system.2012.05.001>
- Borg, S. (2006). *Teacher cognition and language education*. Bloomsbury.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Press syndicate of the University of Cambridge. <https://rm.coe.int/16802fc1bf>
- Council of Europe. (2018). *Common European framework of reference for languages: Learning, teaching, assessment – companion volume with new descriptors*. Press syndicate of the University of Cambridge. <https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4>
- Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11, 227–268. https://doi.org/10.1207/S15327965PLI1104_01

- Dewaele, J. (2020). If classroom emotions were music, teachers would be conductors and learners would be members of the orchestra. In O. Mentz & K. Papaja (Eds.), *Focus on language* (pp. 8–10). Mouton De Gruyter.
- Dewaele, J., Witney, J., Saito, K., & Dewaele, L. (2018). Foreign language enjoyment and anxiety: The effect of teacher and learner variables. *Language Teaching Research*, 22(6), 676–697. <https://doi.org/10.1177/1362168817692161>
- Dönszelmann, S. (2019). *Doeltaal-Leertaal: Didactiek, professionalisering en leereffecten*. PhD Dissertation Vrije Universiteit Amsterdam. Parthenon. <https://research.vu.nl/en/publications/target-language-a-vehicle-for-language-learning-pedagogy-professi> [Target language – learning language: didactics, professional development and learning outcomes.]
- Dönszelmann, S. Van Beuningen, C., Kaal, A., & De Graaff, R. (Eds.). (2020). *Handboek Vreemdetalenonderwijs*. Coutinho. [Handbook for modern language teaching].
- Dörnyei, Z., MacIntyre, P. D., Henry, A., & Al-Hoorie, A. H. (2015). *Motivational dynamics in language learning*. Multilingual Matters.
- Fasoglio, D., & Tuin, D. (2017). *Eindniveau gespreksvaardigheid Engels in het voortgezet onderwijs*. SLO. <https://www.slo.nl/publicaties/@4301/taalprofielen-2015/> [Achieved levels of English fluency in secondary education]
- Fasoglio, D., de Jong, K., Trimbos, B., Tuin, D., & Beeker, A. (2015). *Taalprofielen, herziene versie van taalprofielen 2004*. SLO. <https://www.slo.nl/publicaties/@4597/eindniveau/> [Language profiles, revised edition of language profiles 2004]
- Gibbons, P. (2015). *Scaffolding language, scaffolding learning* (2nd ed.). Heinemann.
- Hollenstein, T. (2013). *State space grids: Depicting dynamics across development*. Springer.
- Johnson, K. E. (2006). The sociocultural turn and its challenges for second language teacher education. *TESOL Quarterly*, 40(1), 235–257. <https://doi.org/10.2307/40264518>
- Lamey, A., Hollenstein, T., Lewis, M. D., & Granic, I. (2004). *GridWare version 1.1* [Computer software].
- Larsen-Freeman, D., & Anderson, M. (2011). *Techniques & principles in language teaching* (3rd ed.). Oxford University Press.
- Larsen-Freeman, D., & Anderson, M. (2015). *Techniques & principles in language teaching* (3rd ed., [8th print]. ed.). Oxford University Press.
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford University Press.
- Lightbown, P. M. (2015). *Focus on content-based language teaching*. OUP.
- Macintyre, P. D., Gregersen, T., & Mercer, S. (2019). Setting an agenda for positive psychology in SLA: Theory, practice, and research. *The Modern Language Journal*, 103(1), 262–274. <https://doi.org/10.1111/modl.12544>
- Maley, A. (2011). Squaring the circle, reconciling materials as constraint with materials as empowerment. In B. Tomlinson (Ed.), *Materials development in language teaching* (2nd ed., pp. 379–402). Cambridge University Press.
- Mercer, S., & Dörnyei, Z. (2020). *Engaging language learners in contemporary classrooms* (1st ed.). Cambridge University Press.
- Murphy, P. K., Wilkinson, I. A. G., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of Educational Psychology*, 101(3), 740–764. <https://doi.org/10.1037/a0015576>
- Pennings, H. J. M., Van Tartwijk, J., Wubbels, T., Claessens, L. C., Van der Want, A. C., & Brekelmans, M. (2014). Real-time teacher-student interactions – A dynamic systems approach. *Teaching and Teacher Education*, 37, 183–193. <https://doi.org/10.1016/j.tate.2013.07.016>
- Piggott, L. (2019). *First meaning, then form: A longitudinal study on the effects of delaying and reducing explicit form-focused instruction for young adolescent EFL learners*. LOT dissertation series. Utrecht University., <https://dspace.library.uu.nl/handle/1874/386767>
- Scrivener, J. (2012). *Classroom management techniques*. Cambridge University Press.

- Smit, N., Dijk, M. V., Bot, K. D., & Lowie, W. (2022). The complex dynamics of adaptive teaching: Observing teacher-student interaction in the language classroom. *International Review of Applied Linguistics in Language Teaching*, 60(1), 23–40. <https://doi.org/10.1515/iral-2021-0023>
- Tammenga-Helmantel, M., & Maijala, M. (2019). The position of grammar in Finnish, Dutch, and global course books for German as a foreign language. *Language Teaching Research*, 23(5), 562–583. <https://doi.org/10.1177/1362168817752542>
- The Douglas Fir Group (2016). A transdisciplinary framework for SLA in a multilingual world. *Modern Language Journal*, 100, 19–47. <https://doi.org/10.1111/modl.12301>
- Thornbury, S. (2011). Language teaching methodology. In J. Simpson (Ed.), *The routledge handbook of applied linguistics*. Routledge Handbooks Online.
- Van Batenburg, E. S. L., Oostdam, R. J., van Gelderen, A. J. S., Fukkink, R. G., & de Jong, N. H. (2018). Evaluating opportunities in Dutch EFL course books for developing pre-vocational learners' oral interactional ability. *Language Teaching Research*. <https://doi.org/10.1177/1362168818804258>
- Walqui, A., & Van Lier, L. (2010). *Scaffolding the academic success of adolescent English language learners*. WestEd.
- Waninge, F., Dörnyei, Z., & De Bot, K. (2014). Motivational dynamics in language learning: Change, stability, and context. *The Modern Language Journal*, 98, 704–723. <https://doi.org/10.1111/modl.12118>

Nienke Smit is an assistant professor at Utrecht University, the Netherlands. She specializes in observational research focusing on interaction and adaptive instruction. She obtained a PhD from the University of Groningen. Her PhD research focused on observing and measuring the dynamics of adaptive teaching and on the role of scaffolding in English as a foreign language lessons in secondary schools in the Netherlands. She developed two classroom observation instruments which might be used by researchers and practitioners. Email: n.smit@uu.nl

Marijn van Dijk is a professor of developmental psychology at the University of Groningen, the Netherlands. Her research concerns the development of children in interaction with their caregivers and/or teachers. The focus is on interaction and development (language, feeding, reasoning) and the dynamics of learning. Most studies deal with variability and stability in change processes and the observation of interactional behavior in naturalistic settings.

Kees de Bot is an emeritus professor at the University of Groningen, the Netherlands and a professor of applied linguistics at the University of Pannonia, Hungary. His research concerns language development, bilingualism and language attrition approached from a complex dynamic systems perspective. He is well-known for his contributions to modelling bilingualism and one of the founders of Content and Language Integrated Teaching pedagogy.

Wander Lowie is a professor of Applied Linguistics at the University of Groningen, the Netherlands, and a Research Associate of the University of the Free State in South Africa. He is associate editor of *The Modern Language Journal*. He specializes in the application of Dynamic Systems Theory to second language development (learning and teaching). He is the chair of the Dutch national committee for promoting researched-based language pedagogy.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 35

How Do Dutch Teachers Implement Differentiation In Primary Mathematics Education?



Emilie J. Prast and Marian Hickendorff

Abstract Adapting education to students' diverse educational needs is widely recognised as an important, but also complex aspect of effective teaching. In this chapter, we provide insight into how Dutch primary school teachers implement differentiation based on students' current mathematics achievement level. We review evidence from four independent samples in which the same teacher self-assessment questionnaire was administered ($N = 907$ teachers in total), supplemented with qualitative data from various perspectives: external observers, students, and teachers. Based on these sources of information, we identify the following general patterns. Teachers generally implement achievement-based differentiation at least to some extent. That is, student achievement is monitored, and efforts are taken to adapt instruction or practice to students' current achievement level. This is often organised using within-class homogeneous achievement groups. While low-achieving students regularly receive additional instruction, specific instruction for high-achieving students is uncommon. Refined, qualitative strategies to diagnose students' individual educational needs and to adapt education to these individual needs are also used relatively infrequently. These relatively infrequently used strategies point to areas for improvement. Furthermore, the flexibility of within-class achievement groups seems to vary and deserves more attention in future research and practice.

Keywords Differentiation · Implementation · Mathematics education · Adaptive teaching · Formative assessment

E. J. Prast (✉) · M. Hickendorff
Institute of Education and Child studies, Leiden University, Leiden, The Netherlands
e-mail: e.j.prast@fsw.leidenuniv.nl

1 Introduction

Adapting education to students' diverse educational needs is widely recognised as an important, but also complex aspect of effective teaching (Kyriakides et al., 2009; Parsons et al., 2018). Implementing differentiation requires specific attitudes, knowledge, and skills, and concerns about suboptimal implementation of differentiation have been raised (Hertberg-Davis, 2009; Inspectorate of Education, 2012, 2018; Schumm et al., 2000; Van Geel et al., 2018; Vogt & Rogalla, 2009). Knowledge about how teachers currently adapt education to students' diverse educational needs is the first step towards promoting effective differentiation. In this chapter, we focus on the research question: How do Dutch primary school teachers implement differentiation based on students' current mathematics achievement level? Specifically, which strategies are used relatively frequently and infrequently? To answer this question, we looked for general patterns in data from four independent studies that investigated differentiation practices in Dutch primary mathematics education using various quantitative and qualitative measures.

1.1 Theoretical Background

In this chapter, we focus on differentiation based on students' current level of knowledge and skills (also called readiness-based or cognitive differentiation), defined as 'an approach by which teaching is varied and adapted to match students' abilities using systematic procedures for academic progress monitoring and data-based decision-making' (Roy et al., 2013, p.1187). According to this definition, teachers should monitor students' academic progress to identify students' educational needs and then adapt instruction to these needs.

To specify *how* educational needs should be determined and *how* instruction should be adapted in the context of primary mathematics education, a previous study sought consensus among experts in the field of differentiation and mathematics education (Prast et al., 2015). This resulted in the cycle of differentiation depicted in Fig. 35.1.

Organisationally, this model assumes the use of flexible homogeneous within-class achievement groups (Tieso, 2003). The term 'achievement grouping' rather than 'ability grouping' is used since students should be grouped flexibly based on their current level of knowledge and skills rather than on (presumably fixed) academic ability. These achievement groups (typically a low-achieving, average-achieving and high-achieving group) should be used part of the time to cater specifically to the educational needs of the different subgroups, besides whole-class activities where possible and individualised adaptations where necessary. Note, however, that the steps of the cycle of differentiation could also be organised in a different way.

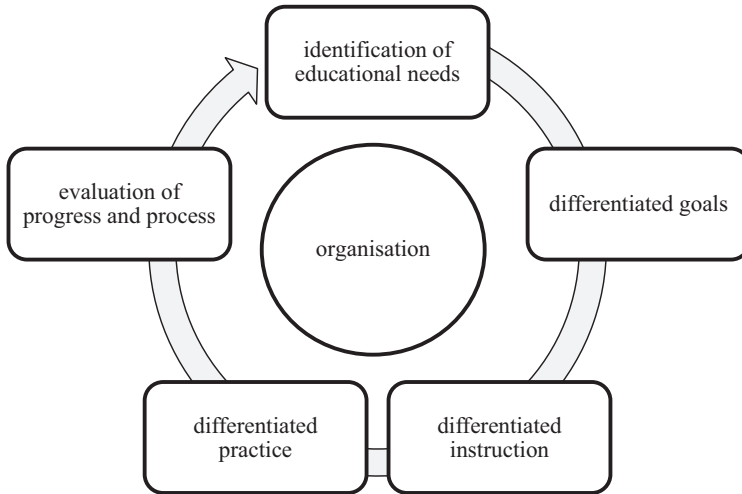


Fig. 35.1 Cycle of differentiation (Prast et al., 2015)

The first step in the cycle of differentiation is the identification of educational needs. Information from various sources, including formal and informal assessments, should be used to assign students to achievement groups, to change these groups when necessary, and to gather more refined information about students' educational needs (Prast et al., 2015; Van Geel et al., 2018). In the second step, the teacher should set challenging but realistic goals, which may be the same (convergent differentiation) or different (divergent differentiation) for the different subgroups (Blok, 2004; Prast et al., 2015; Van Geel et al., 2018). Third, the teacher should differentiate instruction through broad whole-class instruction engaging students of diverse achievement levels, tailored subgroup instruction, and individual adaptations (Bosker et al., 2021; Prast et al., 2015). Effective instructional approaches for low-achieving students in mathematics include direct explicit instruction and adapting the level of abstraction (e.g., starting at a more concrete level by working with manipulatives) (Gersten et al., 2009; Van Groenestijn et al., 2011). High-achieving students may need less instruction to reach the general goals for the whole class, but these students also need instruction and feedback (VanTassel-Baska & Stambaugh, 2005). This may include subgroup instruction that stimulates higher-order thinking and reflection on various possible ways of solving a challenging problem (Prast et al., 2015; Rogers, 2007). Fourth, the practice tasks should be differentiated. For the low-achieving subgroup, the most crucial tasks towards mastery of the goals should be selected. For the high-achieving subgroup, the regular material should be compacted and supplemented with challenging enrichment tasks (Rogers, 2007; VanTassel-Baska & Wood, 2010). Fifth and finally, the teacher should use a range of formal and informal assessments to evaluate whether the students have met the goals and whether the applied adaptations of instruction and practice had the desired effect (Prast et al., 2015). This phase can also be used to

reflect on the learning process with the students (Van Geel et al., 2018). The evaluation phase informs the teacher about students' current achievement level and about instructional approaches that work for these students, completing the cycle and serving as new input for the identification of educational needs.

1.2 *The Dutch Context*

Meelissen et al. (2020) provide a brief overview of the Dutch educational system and the primary mathematics curriculum. Dutch primary school classes typically include students with a broad range of academic ability and achievement levels. To the extent possible, students with special educational needs are included in regular education. Separate special education schools exist for students with more severe problems. Since the enactment of a new law about inclusive education in 2014, regular education teachers perceive an increased need for differentiation (Ledoux et al., 2020).

Traditionally, Dutch students performed well on international comparative studies about mathematics achievement, but the Netherlands are losing their leading position (KNAW, 2009; Mullis et al., 2020). Moreover, while relatively many Dutch students reach at least a basic level of mathematics achievement, relatively few Dutch students perform excellently (Inspectorate of Education, 2021; Meelissen et al., 2020). Concerns about this have spurred the following developments. First, benchmarks (reference levels) have been established to specify what knowledge and skills students should have obtained at the end of primary school (Meelissen et al., 2020). A distinction is made between fundamental goals, which should be reached by 85% of the students, and striving goals, which should be reached by 65% of the students. In the Netherlands, the mathematics curriculum is primarily determined by the textbooks on which teachers rely heavily (Van Zanten & van den Heuvel-Panhuizen, 2017). Most mathematics textbooks have been adapted to work towards these benchmarks, and typically provide differentiated practice tasks at two or three levels. In the last three grades of primary school, the lowest-level tasks prepare students for the fundamental goals rather than the striving goals, which means that students get differentiated opportunities-to-learn. Second, the crucial role of the teacher in promoting students' mathematics achievement has been acknowledged (KNAW, 2009). Third, the government has started to promote data-based decision-making to increase student achievement (Doolaard, 2013a, b; Visscher, 2015). Data-based decision making is closely related to differentiation, especially to its progress monitoring component.

Taken together, these developments have underlined teachers' important role in monitoring students' progress and adapting instruction accordingly. However, the Dutch Inspectorate of Education has expressed concerns that many teachers do not implement differentiation optimally (Inspectorate of Education, 2012, 2018). In this chapter, we investigate how Dutch teachers implement differentiation in primary

mathematics education. Specifically, we aim to identify general patterns of relatively frequently and infrequently used strategies for differentiation across various samples and sources of data.

2 Method

2.1 Overview and Participants

To answer the research questions, we combine data from four independent samples (see Table 35.1 for an overview). In each sample, the Differentiation Self-Assessment Questionnaire (DSAQ; see Sect. 2.2) was administered. Additionally, different types of data (video observations, student reports and additional teacher self-report data) were collected in the individual samples.

Sample 1 (Prast et al., 2015) consisted of 268 teachers of grade 1 through 6 who worked at schools that chose to participate in a large-scale research and professional development project about differentiation. The DSAQ was administered among all teachers at the start of the project. Sample 2 (Prast et al., 2023) consisted of 50 teachers and their students of grade 1, 3 and 5, recruited through the schools at which pre-service teacher training students did their internship. Sample 3 (Van Geel et al., 2022) included 300 teachers recruited through the network of the researchers on social media. Besides teachers of grade 1 through 6, this sample also included 48 Kindergarten teachers (in the Netherlands, two Kindergarten years are integrated in

Table 35.1 Overview of samples and measures

	Sample 1	Sample 2	Sample 3	Sample 4
Publication(s)	Prast et al. (2015), (2018)	Prast et al. (2023)	Van Geel et al. (2022)	Inspectorate of Education (2021)
Year of data collection	2012	2018	2019	2018–2019
Differentiation Self-Assessment Questionnaire (DSAQ)	268 teachers	50 teachers ^a	300 teachers ^a	289 teachers
Additional data	Video observations in a subsample of 55 teachers (the teacher of grade 3 in all participating schools and teachers of grade 1–6 in 6 schools)	Teacher self-reports of grouping practices; Student reports of differentiated activities ($N = 310$)	Teacher self-reports about learning to implement the strategies	A subsample of 110 of the 289 teachers were asked to fill out lesson logbooks; 65 teachers responded.

^a Item-level DSAQ-scores were provided for this book chapter by the authors of the respective publications

primary school before students enter grade 1). Sample 4 (Inspectorate of Education, 2021) was a nationally representative sample of 289 teachers taking part in the Dutch national mathematics assessment 2018–2019. This sample consisted of 228 teachers teaching sixth grade students in regular primary education and 61 teachers teaching students at the end of special primary education. Differences between regular and special primary education teachers in the DSAQ-scores were minimal (Inspectorate of Education, 2021). In each sample most teachers were female (68–94%), which reflects the Dutch population of primary school teachers. Across samples, teachers had an average of 14–16 years of teaching experience, with a broad range from beginning teachers to very experienced teachers (range 0–44 years). Further details regarding the samples can be found in the respective publications.

2.2 Measures

The Differentiation Self-Assessment Questionnaire (DSAQ; Prast et al., 2015) was developed based on the cycle of differentiation described in Sect. 1.2. For each step in the cycle, a subscale was created comprising items representing various strategies for differentiation (e.g., ‘I analyse the answers on curriculum-based tests to assess a student’s educational needs’; see Tables 35.2a, 35.2b, 35.2c, 35.2d, and 35.2e). Teachers evaluate their use of the strategies on a five-point scale ranging from ‘does not apply to me at all’ to ‘fully applies to me’. In the original validation study, which is Sample 1 in the current chapter, the DSAQ demonstrated convergent and divergent validity compared to other teacher self-assessment scales (Prast et al., 2015). The subscales had an adequate internal consistency (Cronbach’s alpha between 0.69 and 0.86; see Tables 35.2a, 35.2b, 35.2c, 35.2d, and 35.2e for Cronbach’s alpha in each sample). Consistent with Roy et al. (2013), the subscales loaded on two

Table 35.2a DSAQ subscale 1 statistics: reliability (Cronbach’s α) and means and standard deviations (in parentheses) of subscale and individual items

DSAQ subscale / item	Sample 1 (N = 268)	Sample 2 (N = 50)	Sample 3 (N = 300)	Sample 4 (N = 289)	Overall across Samples
	$\alpha = .69$	$\alpha = .68$	$\alpha = .63$	$\alpha = .74$	
Subscale 1: Identification of educational needs	3.64 (0.55)	3.97 (0.55)	3.94 (0.62)	3.83 (0.67)	3.82 (0.61)
1.1 I analyse the answers on curriculum-based tests to assess a student’s educational needs	4.02 (0.77)	4.20 (0.887)	4.13 (0.93)	4.15 (0.91)	4.11 (0.88)
1.2 I analyse the answers on standardised tests to assess a student’s educational needs	3.49 (0.91)	4.00 (0.87)	3.98 (1.01)	3.97 (0.97)	3.83 (0.96)
1.3 I assess specific students’ educational needs based on daily maths work	3.75 (0.72)	4.08 (0.72)	3.95 (0.95)	3.87 (0.89)	3.87 (0.86)
1.4 I assess specific students’ educational needs based on (informal) observations during the maths lesson	3.76 (0.77)	4.20 (0.67)	4.14 (0.90)	3.78 (0.93)	3.92 (0.86)
1.5 If necessary, I conduct diagnostic conversations to analyse the educational needs of specific students	3.20 (0.90)	3.35 (1.05)	3.49 (1.21)	3.35 (1.06)	3.92 (1.07)

Scale: 1 = does not apply to me, 5 = fully applies to me

Color coding: dark green = +0.5 SD compared to overall subscale mean, light green = +0.25 SD, light red = -0.25 SD, dark red = -0.5 SD

Table 35.2b DSAQ subscale 2 statistics: reliability (Cronbach's α) and means and standard deviations (in parentheses) of subscale and individual items

DSAQ subscale / item	Sample 1 (N = 268)	Sample 2 (N = 50)	Sample 3 (N = 300)	Sample 4 (N = 289)	Overall across Samples
	$\alpha = .79$	$\alpha = .73$	$\alpha = .73$	$\alpha = .81$	
Subscale 2: Differentiated goals	3.78 (0.55)	3.83 (0.67)	4.05 (0.66)	4.22 (0.67)	3.92 (0.63)
2.1 I set different goals for the children, dependent on their achievement level	3.62 (0.79)	3.67 (1.02)	3.80 (1.03)	4.25 (0.93)	3.88 (0.93)
2.2 I set extra challenging goals for high-achieving students	3.57 (0.83)	3.84 (1.04)	4.21 (0.90)	4.18 (0.98)	3.99 (0.91)
2.3 I set well-considered minimum goals for very low-achieving students	3.75 (0.76)	3.78 (1.00)	3.87 (1.04)	4.20 (0.89)	3.93 (0.91)
2.4 I know the opportunities for differentiation offered by the curriculum	4.03 (0.68)	3.98 (1.00)	4.28 (0.96)	4.33 (0.81)	4.21 (0.84)
2.5 I use the opportunities the curriculum offers for differentiation for high-achieving students	3.88 (0.84)	4.08 (0.93)	3.97 (1.05)	4.33 (0.95)	4.02 (0.95)
2.6 I use the opportunities the curriculum offers for differentiation for low-achieving students	3.83 (0.82)	3.65 (1.11)	3.97 (1.05)	4.19 (0.92)	3.98 (0.95)

Table 35.2c DSAQ subscale 3 statistics: reliability (Cronbach's α) and means and standard deviations (in parentheses) of subscale and individual items

DSAQ subscale / item	Sample 1 (N = 268)	Sample 2 (N = 50)	Sample 3 (N = 300)	Sample 4 (N = 289)	Overall across Samples
	$\alpha = .72$	$\alpha = .70$	$\alpha = .77$	$\alpha = .86$	
Subscale 3: Differentiated instruction	3.81 (0.42)	4.19 (0.44)	4.13 (0.58)	4.21 (0.62)	4.07 (0.55)
3.1 I adapt the level of abstraction of instruction to the needs of the students	3.95 (0.55)	4.48 (0.54)	4.27 (0.76)	4.41 (0.74)	4.23 (0.69)
3.2 I adapt the modality of instruction (visual, verbal, manipulative) to the needs of the students	3.82 (0.62)	4.22 (0.68)	4.17 (0.84)	4.22 (0.78)	4.09 (0.69)
3.3 I adapt the pace of instruction to the needs of the students	3.95 (0.56)	4.34 (0.77)	4.28 (0.81)	4.40 (0.74)	4.22 (0.72)
3.4 I deliberately ask open-ended questions during whole-class instruction	3.82 (0.67)	4.20 (0.76)	4.12 (0.98)	4.24 (0.93)	4.07 (0.87)
3.5 I deliberately ask questions at various difficulty levels during whole-class instruction	3.69 (0.73)	4.16 (0.82)	4.09 (0.96)	4.08 (0.89)	3.97 (0.87)
3.6 I regularly provide low-achieving children with additional instruction (extended instruction, pre-teaching)	4.25 (0.64)	4.42 (0.70)	4.45 (0.81)	4.37 (0.83)	4.36 (0.76)
3.7 I regularly provide high-achieving students with additional instruction or guidance at their level, in a group or individually	3.20 (0.92)	3.51 (0.85)	3.58 (1.13)	3.80 (0.94)	3.53 (1.00)

Table 35.2d DSAQ subscale 4 statistics: reliability (Cronbach's α) and means and standard deviations (in parentheses) of subscale and individual items

DSAQ subscale / item	Sample 1 (N = 268)	Sample 2 (N = 50)	Sample 3 (N = 300)	Sample 4 (N = 289)	Overall across Samples
	$\alpha = .72$	$\alpha = .69$	$\alpha = .80$	$\alpha = .74$	
Subscale 4: Differentiated practice	3.46 (0.55)	3.73 (0.57)	3.83 (0.73)	3.96 (0.70)^b	3.66 (0.64)^c
4.1 I vary different types of practice during the maths lesson (e.g. individual or group work, solution spoken, written or drawn)	3.53 (0.78)	4.02 (0.74)	3.79 (1.04)	3.79 (0.92)	3.70 (0.91) ^c
4.2 I adjust different types of practice to the needs of the students in the classroom (e.g. having a specific child complete exercises on the computer because this child learns more in this way)	3.04 (0.83)	3.31 (0.95)	3.52 (1.13)	3.58 (0.95)	3.29 (1.00) ^c
4.3 I select the most important tasks for very low-achieving students	3.73 (0.73)	3.72 (0.93)	4.03 (1.01)	4.09 (0.91)	3.87 (0.89) ^c
4.4 I use curriculum compacting for high-achieving students	3.20 (1.25)	4.00 (1.14)	3.97 (1.20)	4.07 (1.16)	3.64 (1.22) ^c
4.5 I provide high-achieving students with enrichment tasks	4.00 (0.87)	4.23 (1.04)	4.37 (0.94)	4.27 (1.00)	4.20 (0.92) ^c
4.6 I also use computer programmes or maths websites in my maths lessons	3.68 (0.97)	3.92 (1.03)	3.96 (1.78)	– ^a	3.84 (1.43) ^c
4.7 I use computer programmes and/or maths websites to offer children focused practice in a skill that they do not sufficiently master	3.32 (0.96)	3.46 (1.18)	3.59 (1.31)	– ^a	3.46 (1.16) ^c
4.8 I use computer programmes and/or maths websites to offer specific children additional challenge in the maths lesson	3.15 (1.05)	3.18 (1.17)	3.44 (1.36)	– ^a	3.29 (1.22) ^c

^aThis item was not administered in Sample 4 due to overlap with other items in the questionnaire of that study

^bThe scale mean and standard deviation were computed on items 4.1 through 4.5

^cThe overall means and standard deviations were computed based on Sample 1–3

Table 35.2e DSAQ subscale 5 statistics: reliability (Cronbach’s α) and means and standard deviations (in parentheses) of subscale and individual items

	Sample 1 (N = 268)	Sample 2 (N = 50)	Sample 3 (N = 300)	Sample 4 (N = 289)	Overall across Samples
DSAQ subscale / item	$\alpha = .86$	$\alpha = .80$	$\alpha = .77$	$\alpha = .80$	
Subscale 5: Evaluation of progress and process	3.56 (0.57)	3.78 (0.60)	3.81 (0.63)	3.87 (0.61)	3.75 (0.61)
5.1 use scores on standardised and curriculum-based tests to evaluate whether the learning goals have been met	4.04 (0.73)	4.31 (0.74)	4.24 (0.93)	4.42 (0.82)	4.24 (0.83)
5.2 analyse the answers on curriculum-based tests to evaluate whether the learning goals of that unit have been met	4.06 (0.72)	4.33 (0.83)	4.31 (0.91)	4.22 (0.93)	4.21 (0.86)
5.3 regularly evaluate whether all students have met the learning goals based on their daily maths work	3.75 (0.85)	3.96 (0.86)	4.00 (0.93)	4.05 (0.91)	3.94 (0.90)
5.4 evaluate whether all students have met the lesson goals based on (informal) observations during the maths lesson	3.45 (0.86)	3.58 (0.99)	3.94 (0.94)	3.83 (0.88)	3.74 (0.90)
5.5 conduct diagnostic conversations to evaluate whether specific students have met the lesson goals	2.85 (0.87)	3.10 (1.09)	2.96 (1.11)	3.13 (1.00)	2.99 (1.01)
5.6 evaluate whether the type of instruction and practice chosen by me were effective for the majority of the students in the class	3.44 (0.77)	3.72 (0.81)	3.74 (0.96)	3.78 (0.88)	3.66 (0.87)
5.7 evaluate whether a specific type of instruction was effective for specific students	3.32 (0.80)	3.46 (0.91)	3.54 (1.00)	3.67 (0.86)	3.51 (0.90)

higher-order factors, namely progress monitoring (subscales identification of educational needs and evaluation of progress and process) and instructional adaptations (subscales differentiated goals, differentiated instruction, and differentiated practice).

A brief description of the additional data collected in the individual samples is integrated in the results section to enhance readability.

3 Results

3.1 DSAQ Results

Mean scores and pooled standard deviations across all four samples were calculated. As Tables 35.2a, 35.2b, 35.2c, 35.2d, and 35.2e show, teachers’ self-assessment scores were generally quite high, with mean scores well above the midpoint of the scale for all subscales and for most items.

To investigate which strategies for differentiation had relatively high and low scores, the mean item scores were compared to the mean score for the subscale to which each item belonged, in relation to the pooled standard deviation of the subscale. Specifically, item scores were considered moderately high (light green in Tables 35.2a, 35.2b, 35.2c, 35.2d, and 35.2e) if they were at least a quarter of a standard deviation higher than the subscale mean and high (dark green) if they were at least half of a standard deviation higher than the subscale mean. Similarly, item scores were considered moderately low (light red) if they were at least a quarter of a standard deviation lower than the subscale mean and low (dark red) if they were at least half a standard deviation lower than the subscale mean. This is reported per sample as well as for the overall results aggregated across the samples.

As can be seen in Tables 35.2a, 35.2b, 35.2c, 35.2d, and 35.2e, the pattern of (moderately) low or high use was largely consistent across samples. Strategies that were classified as (moderately) high in the overall sample were not always (moderately) high compared to the subscale average of each individual sample, but they were almost never classified as (moderately) low in individual samples. The same goes for strategies that were classified as (moderately) high. Only for two items (4.4 and 4.6), the direction of effects differed between samples, with moderately low scores in Sample 1 and moderately high scores in Sample 2.

We continue to describe the overall scores across the four samples. Teachers indicated to use various sources of information to identify students' educational needs (range 3.83–4.11), with moderately high scores for the analysis of answers on curriculum-based tests, and low scores for diagnostic conversations. Regarding differentiated goals, item scores were quite homogeneous (range 3.88–4.21), with only one moderately high score for knowing the opportunities of differentiation offered by the curriculum. Within the subscale for differentiated instruction (range 3.53–4.36), there was a remarkable difference between the high score for additional instruction for low-achieving students and the low score for instruction for high-achieving students. Adapting the pace of instruction also scored moderately high. Regarding differentiated practice (range 3.29–4.20), there was substantial variability between the items. While the general use of varied types of practice was around the subscale average, the score for adjusting different types of practice to the needs of specific students was low. Selection of the most important tasks for very low-achieving students scored moderately high, and the use of enrichment tasks for high-achieving students scored high. While the general use of computer programmes was moderately high, the use of computer programmes for focused practice was moderately low and the use of computer programmes for specific challenge was low. Regarding evaluation (range 2.99–4.24), the reported use of scores on standardised and curriculum-based tests to evaluate students' progress was high, and the use of daily mathematics work was moderately high. In contrast, evaluating whether a specific type of instruction was effective for specific students scored moderately low and conducting diagnostic conversations to evaluate whether specific students have met the lesson goals scored low.

3.2 *Additional Data*

In each sample, additional data were collected using various measures. In this section, the most relevant results are summarised.

In a subsample of 55 teachers from Sample 1, one or two mathematics lessons per teacher were observed and scored with a systematic video observation instrument (see Prast et al., 2018, for details). The results indicated that most teachers worked systematically with achievement groups. Most teachers differentiated the practice tasks based on the suggestions in the textbook, sometimes complemented with supplementary materials. For high-achieving students, the use of challenging

enrichment tasks was more common than compacting of the regular material (i.e., reducing the amount of repetitive practice). Regarding instructional attention and adaptations, the observations revealed a difference between differentiation for low-achieving and high-achieving students. Teachers regularly spent specific attention to low-achieving students, for example by providing extended instruction to a subgroup, providing explicit instruction, teaching at a lower level of abstraction, or building understanding of the concepts behind a mathematical procedure (i.e., multiplication and division). In contrast, specific instructional attention for high-achieving students was very seldomly observed. Only a few teachers ever spent more than one minute specifically with high-achieving students across the observed lessons. Some teachers did differentiate instruction for high-achieving students by allowing them to skip the whole-class instruction.

In Sample 2 (Prast et al., 2023), two types of additional data were collected: interviews with teachers about their achievement grouping practices, and student questionnaires about their perceptions of differentiated activities in mathematics lessons. In structured interviews, teachers ($N = 50$) were asked whether and how they used achievement groups. Most teachers indicated to use achievement grouping in some way, either fully integrated in their mathematics teaching routine to differentiate instruction and practice ($n = 32$, 64%) or partly ($n = 14$, 28%), for example using the achievement groups for subgroup instruction but not for differentiation the practice tasks. Of the teachers using achievement groups (partly or fully), most teachers reported to create and update grouping arrangements periodically based on students' achievement on curriculum-based or standardised tests. Specifically, 11 teachers (22%) reported to make new grouping arrangements twice per year, 6 teachers (12%) three to four times per year, and 15 teachers (30%) approximately every three to six weeks based on each curriculum-based test. Some of these teachers indicated that these groups could be adapted per lesson based on students' needs. Another 8 teachers (16%) indicated to work with flexible groups, created per lesson or per week based on teachers' assessment of students' educational needs, on educational software or on students' own view on whether they needed additional instruction. The remaining teachers did not change the groups ($n = 1$, 2%), changed grouping arrangements in a different way ($n = 3$, 6%) or had missing responses ($n = 2$, 4%).

In the student questionnaire, students of the teachers in Sample 2 were asked to rate how often they participated in various differentiated and undifferentiated activities such as whole-class instruction, subgroup instruction and working at more or less difficult tasks (see Prast et al., 2023, for details). The questionnaire was administered in written form among all students with informed consent of grade 3 and 5, and as an individual interview among randomly selected low-achieving, average-achieving and high-achieving students of grade 1. Additionally, scores on a standardised mathematics achievement test were collected. $N = 310$ students (21 students of grade 1, 139 students of grade 3, and 150 students of grade 5) provided data on the questionnaire and on the achievement test. The results indicated that student-reported activities were clearly differentiated by achievement level: low-achieving students more frequently reported to receive extended instruction in a subgroup or individually and to work on less difficult tasks, whereas high-achieving

students more frequently reported to work at enrichment tasks. However, high-achieving students (and students of other achievement levels) rarely reported to receive subgroup instruction or individual instruction about enrichment tasks and reported to work independently significantly more often than lower-achieving students.

In Sample 3 (Van Geel et al., 2022), teachers were asked how much time and effort it took to learn to use each of the differentiation strategies included in the DSAQ. Teachers' self-reported use of the strategies correlated negatively with teachers' perceived time and effort to learn the strategies. In other words, strategies that were considered easy to learn were implemented more frequently. Additionally, teachers were asked about facilitators and barriers for learning to implement the differentiation strategies. Gaining experience and developing (unspecified) attitudes and beliefs were considered the most helpful factors, whereas limited time management and a lack of experience and were considered the most impeding ones. Interestingly, (limited) skills and knowledge gained during initial teacher training were frequently identified as facilitator *and* barrier, perhaps due to differences between teacher training institutes regarding the way in which aspiring teachers learn to differentiate. Finally, teachers with less than three years of experience were shown to score lower on the DSAQ.

In Sample 4 (Inspectorate of Education, 2021), a subsample of 65 teachers kept logbooks of one to four randomly selected mathematics lessons. To identify students' educational needs, teachers most often reported to use students' daily work (55.4% of the teachers used this at least once across the reported lessons), followed by scores on achievement tests (30.1%) and other measures (19.3%). Teachers most frequently used these data to analyse students' mistakes, to assign students to achievement groups, and to determine students' mastery of the content. Approximately one-fifth of the teachers (21.7%) did not use any data to monitor students' progress in the reported lessons. Regarding adaptations, teachers most frequently mentioned to adapt instruction (66.2%), followed by goals (33.1%) and practice (28.3%), although these categories sometimes overlapped. Frequently mentioned adaptations were shortened or extended instruction, working with homogeneous achievement groups, differentiation of the practice tasks (amount or difficulty level) and individual instruction or support. Approximately one-fifth of the teachers (21.1%) did not make any adaptations in goals, instruction, or practice across the reported lessons.

4 Conclusion and Discussion

4.1 General Patterns

The aim of the current study was to chart teachers' differentiation practices in primary mathematics by identifying relatively frequent and infrequent strategies. We integrated the findings of four different studies that had the teacher self-report questionnaire (DSAQ) in common, which was accompanied by additional, more

qualitative data (videos and lesson logs) in two of these studies. We identified several general patterns of relatively frequently and infrequently reported strategies that were similar across samples and measures. The two main components of differentiation – progress monitoring and instructional adaptations (Roy et al., 2013) – are clearly implemented by most teachers at least in a basic way. Most teachers monitor students' achievement using standardised tests, curriculum-based assessments and students' daily work. These assessments are used to identify students' educational needs and frequently form the basis for creating homogeneous within-class achievement groups. Based on this assessment of students' achievement level and educational needs, instructional adaptations are made.

A typical differentiated lesson could look like this. First, the teacher provides a whole-class instruction. Sometimes, high-achieving students already start to work independently during the whole-class instruction. After the whole-class instruction, average-achieving and high-achieving students work independently at tasks provided by the textbook, which are typically differentiated at three levels. Simultaneously, the teacher provides extended instruction to a subgroup of low-achieving students. The extended instruction may be at a slower pace, at a lower level of abstraction, or more explicit than the whole-class instruction. Subsequently, all students continue to work independently, while the teacher monitors and addresses individual questions where necessary. When high-achieving students finish their regular work, they move on to enrichment tasks provided by the textbook or supplementary materials. Finally, the teacher may conclude the lesson with a whole-class wrap-up, in which the teacher reflects with the students on what they have learned.

In contrast to these frequently implemented strategies for differentiation, other strategies were less frequently reported and observed. While teachers routinely provide extended instruction to low-achieving students, teachers infrequently provide specific instruction to high-achieving students (for example, about enrichment tasks), which may signal a tendency for convergent rather than divergent differentiation. Furthermore, some of the more refined, qualitative and individually tailored strategies for differentiation are used relatively infrequently. Specifically, teachers infrequently use diagnostic conversations to gain qualitative information about students' educational needs and infrequently evaluate whether a specific instructional adaptation was effective for individual students. Furthermore, teachers do not frequently adjust the type of practice to students' needs. The use of computer programmes for additional specific practice or challenge was also relatively infrequently reported.

4.2 *Limitations and Strengths*

The following limitations should be considered. Selection bias may have played a role in some of the samples. Especially Sample 1 (teachers at schools that were interested in an extensive professional development programme about

differentiation) and Sample 3 (teachers recruited through social media) may have included teachers with a special interest for differentiation, although this bias could go in two directions: teachers could be interested because they feel the need to improve their differentiation skills, or because they already spend a lot of attention to differentiation. Nevertheless, the pattern of relatively frequently and infrequently reported strategies was similar across samples. Moreover, the combination of the four independent samples is quite large and diverse, representing a variety of schools from multiple regions in the Netherlands, and teachers with various levels of experience.

Another limitation is the use of a teacher self-report questionnaire as the primary measure. Teachers might rate their own use of differentiation differently than external observers. Therefore, we complemented these findings with qualitative findings from different perspectives, namely external observers and students. Although the main patterns of relatively frequently and infrequently used strategies described above were largely consistent across different perspectives, the general level and quality of implementation cannot be directly compared across these measures. More observational studies, in which the quality of implementation can also be examined in more detail, would be desirable in future research.

4.3 Implications for Research and Practice

The finding that many teachers implement basic strategies for differentiation such as monitoring student progress with tests and using differentiated practice tasks provided by the mathematics textbook is in line with previous national and international studies (Inspectorate of Education, 2018; Roy et al., 2013), in which it was found that such strategies are implemented relatively frequently compared to other strategies which require more time or skills to implement. The implementation of these basic strategies for differentiation may have been further supported by the increased attention for data-based decision-making and differentiation in professional development programs, as well as by the extensive suggestions for differentiation in recent versions of mathematics textbooks. At the same time, the differences between teachers should not be overlooked: while most teachers in the current study implemented differentiation at least to some extent, the qualitative findings in Sample 3 also indicated that about 20% of the teachers did not monitor progress and did not adapt goals, instruction, or practice in any way in the reported lessons. Future research might investigate what explains these differences between teachers.

The widespread use of achievement grouping warrants more research about the way in which teachers implement this, in the Netherlands but also in other countries. Specifically, the flexibility of achievement groups should receive more attention in future research and practice. Based on the single study (Prast et al., 2023), in which this topic was examined, it seems that the flexibility of achievement groups differs substantially between teachers. Some teachers used achievement groups flexibly, deciding on a lesson-by-lesson basis which students needed additional

instruction and which practice tasks would be most suitable (sometimes assisted by educational software). In this case, achievement groups are used as a means to adapt instruction and practice to students' current educational needs, as recommended (Prast et al., 2015). However, a substantial percentage of teachers used achievement groups in a less flexible way, updating them for example only twice a year after the administration of a standardised mathematics achievement test. Fixed achievement groups are problematic, because they are less responsive to changes in students' educational needs (which may also vary per topic). Moreover, when students placed in low achievement groups have limited opportunities to move to a higher achievement group, this may limit their future educational chances (Denessen, 2017; Van den Bergh, 2018). While we cannot draw strong conclusions based on the single study described in this chapter, teachers should be aware of the importance of the flexibility of achievement groups and more research into this topic is needed. Substantial differences between countries regarding the use and flexibility of achievement groups may be expected. For example, within-class achievement grouping is commonly used in the Netherlands, while other countries, including the UK, have a tradition of between-class achievement grouping (Hallam & Parsons, 2013). Such organisational factors may affect the flexibility of the achievement groups.

Areas for improvement concern the relatively infrequently used strategies for differentiation. The limited specific instructional attention for high-achieving students is consistent with a previous study and might partly explain the relatively low percentage of excellent-achieving students in the Netherlands compared to other countries (Inspectorate of Education, 2019; Mullis et al., 2020). However, concerns about limited attention for high-achieving or gifted students in general education have also been raised previously by researchers from other countries including the US (Brighton et al., 2015; Hertzberg-Davis, 2009). When high-achieving students work at sufficiently challenging enrichment tasks, they also need instruction or feedback about these tasks (VanTassel-Baska & Stambaugh, 2005). Moreover, differentiation for high-achieving students could generally be more systematic and goal-directed: teachers often provide students with enrichment tasks, but a risk is that these are used to keep students occupied rather than as a means to reach a higher learning goal (Inspectorate of Education, 2019; VanTassel-Baska & Stambaugh, 2005). Another area for improvement concerns refined and qualitative strategies to diagnose students' individual educational needs and adapt instruction and practice to these. This is in line with previous international reviews, although most of the reviewed studies were carried out in the US (McKenna et al., 2015; Scott et al., 1998). While the implementation of such strategies might improve the fit of educational practices to students' individual educational needs, implementing such strategies requires substantial time and effort from the teacher. Therefore, the extent of individual differentiation that is realistic to expect from general education teachers should also be considered.

In all areas for improvement, future research could examine why these strategies are relatively infrequently used and how they could be promoted, for instance in teacher education and professionalisation. Explanatory factors could be teacher

attitudes and beliefs (e.g., a fixed mindset (Dweck, 2006) as an implicit reason for using fixed achievement groups), teacher knowledge and skills (e.g., being able to provide subgroup instruction about enrichment tasks or to hold a diagnostic conversation), or time and resources (e.g., time to provide subgroup instruction to high-achieving students; available instructional materials; support from colleagues). Based on the findings in Sample 3 (Van Geel et al., 2022), each of these factors seems to be relevant. While more experienced teachers reported a higher level of implementation of differentiation, teachers also reported that attitudes, pre-service teacher education and (limited) time were important facilitators or barriers in learning to implement the strategies. In addition, future research could examine the role of the teaching context in the effectivity and suitability of the various strategies. Depending on factors such as class size, heterogeneity of achievement level, and the number of students with special educational needs in a given class, some strategies may be more effective or suitable than others. For example, in a context where most students struggle to reach the basic lesson goals, it might be a valid choice to focus all efforts on reaching these basic goals at the expense of subgroup instruction about enrichment tasks. Thus, while pre-service teacher education and professional development programs for in-service teachers should strive to provide teachers with the necessary attitudes, knowledge and skills to implement differentiation, the importance of taking into account the classroom context and providing teachers with sufficient time and resources for implementation should not be overlooked.

References

- Blok, H. (2004). Adaptief onderwijs: Betekenis en effectiviteit [Adaptive education: A concept analysis and a review of research results]. *Pedagogische Studiën*, 81(1), 5–27.
- Bosker, R. J., Durgut, F., Edzes, H., Jol, M., Van Tuijl, C., & Van der Vegt, A. L. (2021). *Leidraad differentiatie als sleutel voor gelijke kansen [Guideline differentiation as a key to equal opportunities]*. NRO. www.onderwijskennis.nl/leidraad-differentiatie
- Brighton, C. M., Moon, T. R., & Huang, F. H. L. (2015). Advanced readers in reading first classrooms: Who was really “Left behind”? Considerations for the field of gifted education. *Journal for the Education of the Gifted*, 38(3), 257–293. <https://doi.org/10.1177/0162353215592501>
- Denessen, E. (2017). *Verantwoord omgaan met verschillen: Sociale-culturele achtergronden en differentiatie in het onderwijs [Dealing with differences responsibly: Social-cultural backgrounds and differentiation in education]*. Leiden University. <https://hdl.handle.net/1887/51574>
- Doolaard, S. (2013a). *Effecten van het trainings- en begeleidingstraject “Streef: Gebruik maken van opbrengsten” [Effects of the professional development programme “Strive: using achievement data”]*. GION/RUG. <https://pure.rug.nl/ws/portalfiles/portal/2386147/Effectentraining.pdf>
- Doolaard, S. (2013b). *Het streven naar kwaliteit in scholen voor primair onderwijs [Striving for quality in primary schools]*. GION/RUG. <https://core.ac.uk/download/pdf/148191112.pdf>
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.
- Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, 79(3), 1202–1242. <https://doi.org/10.1108/JKM-08-2016-0350>

- Hallam, S., & Parsons, S. (2013). The incidence and make up of ability grouped sets in the UK primary school. *Research Papers in Education*, 28(4), 393–420. <https://doi.org/10.1080/0267152.2012.729079>
- Hertberg-Davis, H. (2009). Myth 7: Differentiation in the regular classroom is equivalent to gifted programs and is sufficient: Classroom teachers have the time, the skill, and the will to differentiate adequately. *Gifted Child Quarterly*, 53(4), 251–253. <https://doi.org/10.1177/0016986209346927>
- Inspectorate of Education. (2012). *De Staat van het Onderwijs: Onderwijsverslag 2010–2011 [The state of education: Educational report 2010–2011]*. Inspectie van het Onderwijs. <https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/rapporten/2013/04/01/staat-van-het-onderwijs-2011-2012-kopie/onderwijsverslag-2011-2012-printversie.pdf>
- Inspectorate of Education. (2018). *De Staat van het Onderwijs: Onderwijsverslag 2016–2017 [The state of education: Educational report 2016–2017]*. Inspectie van het Onderwijs. https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/rapporten/2018/04/11/rapport-de-staat-van-het-onderwijs/108126_IvhO_StaatvanhetOnderwijs_TG.pdf
- Inspectorate of Education. (2019). *Reken-en wiskundeonderwijs aan potentieel hoogpresterende leerlingen [Mathematics education for potentially high-achieving students]*. Inspectie van het Onderwijs. <https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/themaraapporten/2019/04/10/themaonderzoek-reken--en-wiskundeonderwijs-aan-potentieel-hoogpresterende-leerlingen/Rapport+Reken--en+wiskundeonderwijs+aan+%28potentieel%29+hoogpresterende+leerlingen.pdf>
- Inspectorate of Education. (2021). *Peil rekenen-wiskunde einde (speciaal) basisonderwijs 2018–2019 [Mathematics at the end of (special) primary education 2018–2019]*. Inspectie van het Onderwijs. <https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/themaraapporten/2021/04/09/peil.rekenen-wiskunde-einde-s-bo-2018-2019/Peil.Rekenen-Wiskunde+einde+%28s%29bo+2018-2019.pdf>
- KNAW. (2009). *Rekenonderwijs op de basisschool [Mathematics education in primary school]*. KNAW. <https://storage.knaw.nl/2022-07/Rekenonderwijs-op-de-basisschool.pdf>
- Kyriakides, L., Creemers, B. P. M., & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education*, 25(1), 12–23. <https://doi.org/10.1016/j.tate.2008.06.001>
- Ledoux, G., Waslander, S., & Eimers, T. (2020). *Evaluatie Passend Onderwijs, eindrapport [Evaluation inclusive education, final report]*. Kohnstamm Instituut. <https://evaluatiepassen-donderwijs.nl/publicaties/eindrapport-evaluatie-passend-onderwijs/>
- McKenna, J. W., Shin, M., & Ciullo, S. (2015). Evaluating reading and mathematics instruction for students with learning disabilities. *Learning Disability Quarterly*, 38(4), 195–207. <https://doi.org/10.1177/0731948714564576>
- Meelissen, M., Hamhuis, E., & Weijn, L. (2020). Netherlands. In D. L. Kelly, V. A. S. Centurino, M. O. Martin, & I. V. S. Mullis (Eds.), *TIMSS 2019 Encyclopedia: Education policy and curriculum in mathematics and science* (pp. 1–9). TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2019/encyclopedia/netherlands.html>
- Mullis, I. V., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. TIMSS & PIRLS International Study Center. <https://timss2019.org/reports>
- Parsons, S. A., Vaughn, M., Scales, R. Q., Gallagher, M. A., Parsons, A. W., Davis, S. G., Pierczynski, M., & Allen, M. (2018). Teachers' instructional adaptations: A research synthesis. *Review of Educational Research*, 88(2), 205–242. <https://doi.org/10.3102/0034654317743198>
- Prast, E. J., Van de Weijer-Bergsma, E., Kroesbergen, E. H., & Van Luit, J. E. H. (2015). Readiness-based differentiation in primary school mathematics: Expert recommendations and teacher self-assessment. *Frontline Learning Research*, 3(2), 90–116. <https://doi.org/10.14768/flr.v3i2.163>
- Prast, E. J., Van de Weijer-Bergsma, E., Kroesbergen, E. H., & Van Luit, J. E. H. (2018). Teaching students with diverse achievement levels: Observed implementation of differentiation in primary mathematics education. In E. J. Prast (Ed.), *Differentiation in primary mathematics education*. Utrecht University.

- Prast, E. J., Stroet, K., Koornneef, A., & Wilderjans, T. F. (2023). What do students think about differentiation and within-class achievement grouping? *Frontline Learning Research*, *11*, 57–93. <https://doi.org/10.14786/flr.v11i1.1079>
- Rogers, K. B. (2007). Lessons learned about educating the gifted and talented: A synthesis of the research on educational practice. *Gifted Child Quarterly*, *51*(4), 382–396. <https://doi.org/10.1177/0016986207306324>
- Roy, A., Guay, F., & Valois, P. (2013). Teaching to address diverse learning needs: Development and validation of a differentiated instruction scale. *International Journal of Inclusive Education*, *17*(11), 1186–1204. <https://doi.org/10.1080/13603116.2012.743604>
- Schumm, J. S., Moody, S. W., & Vaughn, S. (2000). Grouping for reading instruction: Does one size fit all? *Journal of Learning Disabilities*, *33*, 477–488. <https://doi.org/10.3102/074193259801900205>
- Scott, B. J., Vitale, M. R., & Masten, W. G. (1998). Implementing instructional adaptations for students with disabilities in inclusive classrooms a literature review. *Remedial and Special Education*, *19*(2), 106–119. <https://doi.org/10.1177/074193259801900205>
- Tieso, C. L. (2003). Ability grouping is not just tracking anymore. *Roeper Review*, *26*(1), 29–36. <https://doi.org/10.1080/02783190309554236>
- Van den Bergh, L. (2018). *Waarderen van diversiteit in het onderwijs [Valuing diversity in education]*. Fontys Opleidingscentrum Speciale Onderwijszorg. <https://nivoz.nl/uploads/PDF-uploads/Lectorale%20rede%20Van%20den%20Bergh%202018.pdf>
- Van Geel, M., Keuning, T., Frèrejean, J., Dolmans, D., Van Merriënboer, J., & Visscher, A. J. (2018). Capturing the complexity of differentiated instruction. *School Effectiveness and School Improvement*, *30*(1), 51–67. <https://doi.org/10.1080/09243453.2018.1539013>
- Van Geel, M., Keuning, T., & Safar, II. (2022). How teacher develop skills for implementing differentiated instruction: Helpful and hindering factors. *Teaching and Teacher education: Leadership and Professional Development (1)*, 1–11. <https://doi.org/10.1016/j.tatelp.2022.100007>
- Van Groenestijn, M., Borghouts, C., & Janssen, C. (2011). *Protocol ernstige rekenwiskundeproblemen en dyscalculie [Protocol severe mathematics difficulties and dyscalculia]*. Van Gorcum. https://erwd.nl/_downloads/protocol-ernstige-reken-wiskundeproblemen-en-dyscalculie/basisonderwijs/protocol-erwd-po-bso-so.pdf
- Van Zanten, M., & van den Heuvel-Panhuizen, M. (2017). Primary school mathematics in the Netherlands: The perspective of the curriculum documents. In D. Thompson, C. Suurtamm, & M. Huntley (Eds.), *International perspectives on mathematics curriculum*. Information Age Publishing.
- VanTassel-Baska, J., & Stambaugh, T. (2005). Challenges and possibilities for serving gifted learners in the regular classroom. *Theory into Practice*, *44*, 211–217. https://doi.org/10.1207/s15430421tip4403_5
- VanTassel-Baska, J., & Wood, S. (2010). The Integrated Curriculum Model (ICM). *Learning and Individual Differences*, *20*(4), 345–357. <https://doi.org/10.1016/j.lindif.2009.12.006>
- Visscher, A. J. (2015). *Over de zin van opbrengstgericht(er) werken in het onderwijs [About the value of (more) data-based decision making in education]*. GION. <https://pure.rug.nl/ws/portalfiles/portal/30567909/oratie.pdf>
- Vogt, F., & Rogalla, M. (2009). Developing adaptive teaching competency through coaching. *Teaching and Teacher Education*, *25*(8), 1051–1060. <https://doi.org/10.1016/j.tate.2009.04.002>

Emilie J. Prast is Assistant Professor in Educational Sciences at Leiden University. Her research concentrates on differentiation based on diversity in students' current achievement level. She approaches this topic from various angles, including definition and measurement, implementation, teacher professional development, and effects of differentiation on student motivation and achievement. She aims to strengthen connections between research and educational practice, particularly when teaching students in pre-service teacher education. She has published her work in leading journals including *Learning and Instruction* and *Contemporary Educational Psychology*. email: e.j.prast@fsw.leidenuniv.nl

Marian Hickendorff is Associate Professor in Educational Sciences at Leiden University. Her ambition is to give empirical basis to questions and discussions about primary school mathematics education, such as: What is the effect of presenting a mathematics problem as a story problem? How do children solve problems like 812–784? How well do students at the end of primary school perform? Furthermore, she aims to encourage the application of modern statistical techniques in learning research. She has published her work in leading journals including *Learning and Instruction*, *Journal of Educational Psychology*, and *Learning and Individual Differences*.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 36

Differentiation and Students with Special Educational Needs: Teachers' Intentions and Classroom Interactions



Elisa Kupers, Anke de Boer, Judith Loopers, Alianne Bakker,
and Alexander Minnaert

Abstract Differentiation is mainly linked to differences in learning capacities, but students differ in more domains: differences in motivation, behavior and special educational needs (SEN) are equally relevant. In line with the world-wide trend towards inclusive education, the aim of this chapter is to shed light on Dutch teachers' intentions to differentiate, as well as possible differences in interactions between teachers and students with and without SEN in regular secondary vocational educational education. We first analyzed teachers' online diary entries with regards to their intended differentiation practices for the next lesson. We coded what kind of intentions arise, the level of detail and quality of these intentions and to what kind of differentiation is referred (only cognitive, or possibly also differentiation on domains of behavior, motivation, or students with SEN). Second, we focused on one-to-one classroom interactions between teachers and students with and without special educational needs. We analyzed to what extent there are differences between the interactions of students with and without SEN in terms of teachers' need-supportive teaching and students' engagement. Together, these studies contribute to our understanding of differentiation intentions and practices with regards to meeting the needs of all students in diverse classrooms.

Keywords Differentiation · Special educational needs · Intentions and classroom interactions

E. Kupers (✉) · A. de Boer · J. Loopers · A. Minnaert
Department of Inclusive and Special Needs Education, University of Groningen, Groningen,
The Netherlands
e-mail: w.e.kupers@rug.nl

A. Bakker
Department of Teacher Education, NHL Stenden University of Applied Sciences, Emmen,
The Netherlands

1 Introduction

A worldwide educational trend is that towards more inclusive education of students with special educational needs (SEN) (such as learning difficulties or behavioral problems) into regular schools, resulting in classrooms being more diverse in terms of students' educational needs (De Boer & Kuijper, 2021). In 2014, the *Wet op Passend Onderwijs* (The Dutch Law on Tailored Education) was implemented in the Netherlands. The aim of the law was to guarantee appropriate education for all students, regardless of their SEN. Although special education still exists in the Netherlands, there is a continuous striving towards including more students with SEN in regular education, with extra support allocated on the school level (Ledoux & Waslander, 2020). This increased diversity has gone hand in hand with an expectation of teachers to be aware of these differences and able to adapt their teaching to the individual needs of learners. Indeed, the ability to differentiate teaching has been named as one of the key characteristics of high quality, effective education (Hamre & Pianta, 2005; Deunk et al., 2018; Tomlinson & Imbeau, 2010).

Differentiation at its core is (pro-actively planned) adaption of education to the diverse needs of students (Van Geel et al., 2019; Smale-Jacobse et al., 2019). According to Deunk et al. (2018), differentiation comprises both a careful monitoring of the students' progress and adapting instruction to differences in these levels of progress. The emphasis in this definition is on the (cognitive) levels of the students. Tomlinson defines differentiation in a broader sense, as "an approach to teaching in which teachers proactively modify curricula, teaching methods, resources, learning activities, and student products to address the diverse needs of individual students and small groups of students to maximize the learning opportunity for each student in a classroom" (Tomlinson et al., 2003, p. 120). The 'needs of students' can relate to the level of skill or understanding, but also to differences in interest or learning profiles.

Differentiation practices can take different forms in the classroom. The first step is usually monitoring progress and assessing the needs of the students in preparation of the lesson (Keuning & Van Geel, 2021; Roy et al., 2013). Consequently, teachers can differentiate in content (offering different sources of information and assignments of varying level of difficulty) or in the learning process (by offering additional or different support to some students). Additionally, teachers can differentiate in the end product (by allowing the students to work on different kinds of end products to assess progress on learning goals) or in shaping the learning environment (by providing quiet space for students to work independently, and simultaneously offer space for group work (Tomlinson et al., 2003; Tomlinson & Imbeau, 2010).

Although differentiation is viewed as an essential component of effective teaching, it has also proven to be a notoriously difficult skill for teachers (Van de Grift et al., 2014). This might be because beginning teachers first need to master more basic teaching skills like general effective instruction, classroom management and so on, before this effective instruction can be tailored to the needs of individual students. A challenge in this aspect is that teachers need to attend to the needs of

many students at the same time. Carefully adapted instruction to one student might be detrimental to the other students if the rest of the class is neglected for too long (van de Pol et al., 2015). This might explain why differentiation does not always lead to positive student outcomes (Deunk et al., 2018); differentiation that is not carefully planned and grounded in other dimensions of effective teaching, will not obtain effect.

Because differentiation has proven to be one of the most complex skills for teachers, it requires the teachers to *proactively plan* instruction in response to differences in student levels of readiness, interests and learning profile (Tomlinson & Imbeau, 2010). These authors also know from experience, however, that 'very few' teachers take differentiation into account when planning their lessons (Tomlinson & Imbeau, 2010). Teachers' intentions to differentiate matter because they have proven to be an important prerequisite for teachers' actual inclusive practices in the classroom (Yan & Sin, 2014), although these practices are usually assessed through self-reports rather than observed behavior (Opoku et al., 2020).

As stated before students differ in more than just their cognitive level. This means that "*differentiation according to students' educational needs*" can refer to many different things. A framework for understanding of the (special) educational needs of students can be found in the self-determination theory (Ryan & Deci, 2000). Students have, according to this theory, three basic psychological needs: autonomy, competence and relatedness. The need for autonomy refers to the student's need to be an active agent in shaping one's own learning process and to have a sense of control and choice in the learning environment. Teachers can facilitate the student's feelings of autonomy by providing autonomy support which entails showing respects towards students, fostering relevance and providing the students with meaningful choices (Stroet et al., 2013). The need for competence entails the feeling of being able to attain goals that are personally relevant for students. Teachers can support this by providing structure and adapting their instruction to the student's level of understanding. This strategy closely aligns with adapted, differentiated instruction. Concluding, the need for relatedness refers to the need to have meaningful relationships with both peers in the classroom and with the teacher. Teachers can play an important role here by showing involvement with their students, by dedicating time and resources to the student, and by showing respect and personal interest in their students (Stroet et al., 2013). In sum, self-determination theory can help us better understand what needs are relevant for students, and consequently how differentiated instruction can attend to differences in those needs.

Looking through the lens of self-determination theory, the position of students with SEN in regular education is a vulnerable one. Students with special educational needs (both behavioral as well as learning problems) are relatively often socially neglected or rejected in the classroom (Rademaker et al., 2020; Majorano et al., 2017). Furthermore, teachers report less feelings of closeness and more conflicts with students with challenging behavior, which in the long run can undermine students' need for relatedness (Zee et al., 2017). Regarding the need for autonomy, although the teacher-student relationship might be conflictuous for students with behavioral problems, these students, too, benefit from an autonomy supportive

learning climate (Savard et al., 2013). And finally, regarding the need for competence, especially students with learning problems are at risk for experiencing lower levels of self-efficacy at school (Burden, 2008; Majorano et al., 2017). This raises the important question to what extent teachers are able to fully meet the needs of learners with special educational needs, and makes an exploration of teachers' differentiation skills and practices all the more relevant.

The necessity of differentiation as a component of effective teaching is widely acknowledged, yet teachers seem to struggle to meet the needs of all of their students, especially students with special educational needs. Many studies in the field of (inclusive) education focus on general attitudes towards inclusive education (Van Mieghem et al., 2020) and differentiation (Schwab, 2018). Yet, to increase our understanding of the complexity of differentiation we need to move beyond this and zoom in on what is happening in teachers' lesson-to-lesson intentions and practices. The aim of this chapter is twofold. First, we aim to better understand teacher's intentions to differentiate in each lesson and how these intentions relate to other teacher skills. Second, we aim to zoom in on moment-to-moment interactions between teachers and individual students in the classroom, in order to test whether teachers are able to differentiate according to the three basic psychological needs of students with and without special educational needs. Our research questions are as follows:

1. *What are teachers' intentions for their upcoming lesson, and to what extent do these relate to differentiation practices? (study 1)*
2. *To what extent are there differences in teacher's degree of need-supportive teaching in individual interactions with students with and without special educational needs (SEN)? (study 2)*

2 General Method

2.1 Design

Within the project 'Differentiation Inside Out', fourteen secondary school teachers and 230 students were followed in an intensive longitudinal, observational design for the duration of one school year. Differentiation intentions, practices and efficacy were assessed through interviews, short Ecological Sampling Method (ESM, ecological momentary assessment) questionnaires and lesson observations. Student outcomes (relating to motivation and basic psychological needs) were assessed similarly through ESM questionnaires relating to specific lessons. The Ethical Committee of the department of Educational and Pedagogical Sciences (University of Groningen) approved of the study design and procedures (October 2017). In order to answer the research questions, we describe two studies that were part of this larger project. The first study focuses on the lesson-specific intentions of teachers as described in the ESM questionnaires. The second study zooms in on one-in-one

teacher-student interactions of students with and without SEN that took place in the video-recorded lessons.

3 Study 1: Lesson-Specific Intentions of Teachers

3.1 Method

3.1.1 Participants

In study 1, fourteen teachers who taught second year pre-vocational education (in Dutch: vmbo-gtl/mavo) in regular secondary education on eight different schools throughout the Netherlands participated. The teachers taught either mathematics ($n = 3$), English ($n = 2$) or Dutch (first language) ($n = 9$). Teachers were on average 35.4 years old ($SD = 9.1$). Their teaching experience ranged from less than 5 years to more than 20 years. Prior to the start of the study, the teachers were informed on the aim and procedures of the study and signed an informed consent form.

3.1.2 Procedure and Instruments

All teachers participated with one (in one case two) of their classes in the study for approximately 20 consecutive weeks during one school year, starting between the end of October and early December. The teachers were interviewed and participated in three waves of classroom observations (see Study 2). They were also asked to complete two to four short ESM questionnaires per week via the web platform *u-can-act* (Blaauw et al., 2019), resulting in a maximum of 40–60 repeated measurements per teacher. Compared to questionnaires which measure teachers' intentions 'in general', the advantage of ESM questionnaires are an elimination of recall bias, and a better understanding of the situated and changing nature of teachers' intentions (see Shiffman et al., 2008). At the end of the data collection period, the teachers received a small incentive in the form of a gift certificate, which is common for participants involved in intensive data collections. At the end of each lesson they taught the class with whom they participated, the teachers automatically received a text message on their phone with a personal link to their diary questionnaire. After 12 closed questions on teachers' perception of their own need-supportive teaching during the lesson and their self-efficacy relating to differentiation, the teachers were asked two concluding open questions. First, their intentions for the last lesson they taught was repeated from their previous diary entry, and teachers were asked to what extent they had realized their intentions. Second, teachers were asked for their intentions for the next lesson that they were going to teach this particular class. They could type their answer in a text box. For the purpose of this study, the answers to these last two questions were analyzed.

3.1.3 Analysis

The answers teachers gave about their intentions for the next lesson were analyzed using a combination of closed and open coding, which allows us to account for the richness of the qualitative data (Flick, 2009) while also ensuring a link with the literature on effective teaching. As a first step, we coded all intentions on the domains of the ICALT (Van de Grift, 2007) which measures different domains of effective teaching. The ICALT is based on an empirically derived hierarchy of teaching skills and comprises on the one hand more basic skills such as fostering a positive classroom climate and providing effective instruction for all students, and on the other hand the more complex skills of ‘teaching learning’ to students, and differentiation. In case the teachers’ answers could not be fitted into one of the ICALT domains, new codes were added. The second step was to further analyze the intentions that referred to differentiation. We coded teachers’ intentions with regards to differentiation based on the ways in which was differentiated (based on Tomlinson et al.’s (2003) distinction between content, process, product or learning environment) and on the student characteristics that were mentioned in response to which the differentiation took place (differentiation based on level/pace of students, on interest, or on learning profile (including behavior). Similarly, there was room for adding additional codes to these main categories through open coding. The coding was performed by the first author; in case of doubt, the codes were discussed with the second author. The codes were further analyzed descriptively.

3.2 Results

In total, the 14 teachers filled out 477 diary questionnaires. Because some entries contained more than one intentions, 551 codes were assigned. In the first step, we analyzed to which teaching domain of teaching behaviour (ICALT, Van de Grift et al., 2014) the intentions referred. In addition to the domains included in the ICALT, we found another type of intention in addition: the intention to motivate students (for instance by making the content appealing to them). Of the 551 intentions, 121 referred to differentiation. These differentiation intentions were further analyzed in step 2.

3.2.1 Teachers’ Intentions in Relation to the ICALT Domains

As we can see in Table 36.1, 23.6% of all teachers’ intentions were coded as related to differentiation. The most prominent were intentions relating to instructions (34.5%) such as giving informative feedback or clearly stating lessons goals. Teachers also formulated intentions for more ‘basic’ teaching skills like classroom organization (11.5%) or providing a positive classroom climate (4.9%). Interestingly,

Table 36.1 Examples from the data and number of intentions per domain (percentages between brackets)

Domain	Number of intentions (%)	Examples
No intention	53 (10.3)	<i>"Nothing special"</i> <i>"I don't know, that is after the vacation, I'll see then."</i>
Climate	25 (4.9)	<i>"I want to be a bit more positive."</i> <i>"Keep the calm."</i>
Instruction	177 (34.5)	<i>"Try to make the lesson goals clearer."</i> <i>"The next lesson [...] is on grammar. The class is struggling with this and I hope to provide more clarity on this subject by providing many examples."</i>
Organization	59 (11.5)	<i>"Make a planning for the last period."</i> <i>"Offering structure."</i>
Activating students	26 (5.1)	<i>"I hope to make some time next lesson for activating methods."</i> <i>"Tomorrow I'm going to do an escape room on reading skills. I hope to achieve that they will discuss and work together."</i>
Teaching learning	12 (2.3)	<i>"I am going to let them apply the theory they have learned [...]."</i> <i>"Sharing reading strategies with one another."</i>
Differentiation	121 (23.6)	<i>"Differentiating more in processing the theory."</i> <i>"We are going to repeat the content of ch. 2. [The students] who don't have questions can practice, the students who have questions I will help individually or in small groups."</i>
Motivating students	6 (1.1)	<i>"I want to put the content in a context that is more fun, in order to make a connection to the students experiences."</i> <i>"Stimulating and motivating the students for the content."</i>
Other	72 (14.0)	
Total	551 (100)	

in about one in ten diary entries (10.3%), teachers indicated to have no specific intentions for the next lesson.

3.2.2 Description of Teachers' Differentiation Intentions

In Table 36.2, we further specified the differentiation intentions of the teachers by coding in which *classroom* the differentiation took place: content, process, product or learning environment *element* (Tomlinson & Imbeau, 2010). By far, most (71%) differentiation intentions had to do with differentiating in the learning process. Teachers for instance described how they intended to give weaker students additional instruction while stronger students could work more independently, or to offer instruction on different levels.

In addition to specifying the classroom element, we also analyzed which student characteristics the teacher considered in their intended differentiation (differences in student levels, interests or learning profiles). Most intentions (67.7%) referred to

Table 36.2 Differentiation intentions labeled by classroom element

Classroom element	Number of intentions (%)	Example
Content	19 (15.7)	<i>“Students can choose between different assignments [...]”</i> <i>“Different options for assignments: more challenging for the stronger students [...]”</i>
Process	86 (71.1)	<i>“Extra explanation when the rest are working independently.”</i> <i>“Instructions on different levels.”</i>
Product	0 (0)	
Learning environment	1 (.8)	<i>“I let several students work in the hallway. This made the classroom quieter which caused the students to be more focused on the task.”</i>
Not specified	15 (12.4)	<i>“Try to differentiate more.”</i>
Total	121 (100)	

differentiation for students of different levels of understanding (for instance, providing assignments or instructions on different levels, offering extra help when weaker students needed it). Only 7 intentions (5.7%) referred to differences in student interest or learning profile (for instance, by letting students choose between reading their own novel in class or picking one from the school library). In the other intentions, the student characteristic was not specified (26.4%).

3.3 Discussion

This study provided a unique insight into teachers’ short-term intentions regarding their teaching and differentiation practices. Several things stood out from our data. First and foremost, differentiation as such was relatively rare in teachers’ intentions (only mentioned in 23.6% of cases). Teachers more often formulated intentions relating to more basic teaching skills such as providing overall good quality instruction, creating a positive classroom climate and classroom management. As Van de Grift et al. (2014) remarked, there is an observable hierarchy in the complexity of teaching skills, and teachers’ intentions may reflect differences in skill levels between teachers. Teachers who are preoccupied with more basic aims might have less cognitive space to pro-actively plan for differentiated instruction.

Looking more in depth at teachers’ differentiation intentions, one result was that these intentions are often formulated briefly and in very general terms. This might have had to do with the method of data collection (a brief questionnaire), but it might also be a reflection of their actual intentions. The latter case would be worrisome, as we know from the literature that differentiation is a complex skill that requires pro-active planning (Tomlinson & Imbeau, 2010). Also, detailed and specific behavioral intentions more often lead to actual behavior than vague and non-specific plans (Osch et al., 2010).

4 Study 2: Differentiated One-on-One Interactions Between Teachers and Students with and Without SEN

4.1 Method

4.1.1 Participants

From the fourteen teachers described under Study 1, we selected a subsample of seven teachers for a detailed analysis of video-recorded individual teacher-student interactions. These teachers all chose one of their classes (second year pre-vocational education (in Dutch: vmbo-gtl) to participate in the study for the duration of one school year. The students in these classes were all asked to participate, resulting in a sample of $n = 166$ (43.98% male). In addition, their parents were also asked for informed consent.

4.1.2 Procedure and Instruments

During the school year, three waves of data collection took place at the beginning, middle and end of the school year. For this study, only the data of the first wave are presented. The teachers were asked to conduct their lessons as they normally would. The lessons were filmed with one camera at the back of the classroom, one camera at the front of the classroom and one small wearable camera that could be attached to the teacher's clothing. Because of the focus on individual interactions between teachers and students, only the segments that contained interactions between the teacher and either a single student or a small group of students were transcribed and coded. An interaction begins with the teacher addressing one particular student, or the student making contact with the teacher, for instance by asking a question. The interaction ends with the teacher walking away or addressing another student. The interactions lasted anywhere between a few seconds to several minutes.

Each interaction was coded on the three dimensions of **need-supportive teaching**: autonomy support, structure and involvement on a Likert scale ranging from -3 to 3 with a coding scheme based on Stroet (2014). Below in Table 36.3, examples of behavior on the negative and positive side of each scale are summarized. After training, inter-observer agreement was established on 5 complete lessons (437 interactions). The levels of agreement (intra-class correlations between observers) were 0.736 for autonomy support, 0.677 for structure and 0.808 for involvement, indicating moderate to good levels of agreement.

Special educational needs were assessed from the perspective of the teacher. Teachers were asked to indicate for each student whether students were perceived as having special educational needs, and if so, what the nature of the special educational needs were. These descriptions were afterwards classified in three main categories: behavioral problems, learning problems, or 'other' problems (e.g. a physical disability). With a map of the classroom, the teachers also indicated which student

Table 36.3 Coding scheme need supportive teaching

Autonomy support	Autonomy thwart
Choice	Control
Fostering relevance	Forcing meaningless activities
Showing respect	Showing disrespect
Provision of structure	Chaos
Clarity	No clarity
Guidance	No guidance
Encouragement	Discouragement
Informational feedback	Evaluative feedback
Involvement	Disaffection or rejection
Affection	Disaffection
Attunement	No attunement
Dedication of resources	No dedication of resources
Dependability	No dependability

Based on Stroet (2014)

sat where. In this way, the interaction data could be coupled to the SEN data. The researchers who coded need-supportive teaching were not aware of the presence or absence of special educational needs of the students on the video.

4.1.3 Analyses

Because of the nested structure of the data (interactions are situated in lessons, which are situated in classes/teachers) we performed multilevel analyses. After a check of the assumptions, we estimated multilevel regression models with SEN (recoded as dummy variables) as the explanatory variable, and the three dimensions of need-supportive teaching as outcome variables (one dependent variable per model).

5 Results

5.1 Descriptive Statistics

In total, 2302 one-on-one teacher-student interactions were coded. Of these interactions, 26% (598 interactions) occurred between a teacher and a student with some form of SEN. Looking at behavioral problems and learning problems separately, 16.9% of all interactions that took place were between a teacher and a student with a behavioral problem, while 11.1% of all interactions were between a teacher and a student with a learning problem (note that these percentages do not add up to 26% because students can also have both a learning problem as well as a behavioral

Table 36.4 Descriptive statistics of dependent variables

	Minimum	Maximum	Mean	St. dev.
Autonomy support	-3	3	0.04	0.96
Structure	-3	3	0.45	0.81
Involvement	-3	3	0.43	1.00
Total need-supportive teaching	-6	7	1.21	2.01

problem). Table 36.4 lists the descriptive statistics for the four dependent variables. All variables showed an approximate normal distribution.

5.2 Differences in Teacher-Student Interactions Between Students With and Without SEN

Figures 36.1 and 36.2 show the differences in need-supportive teaching between interactions with students with and without SEN (learning problems and behavioral, respectively). We tested the relation between either two forms of SEN and the three dimensions of need-supportive teaching with multilevel regression models. Although the data has a three-level structure (interactions within students within teachers), exploratory analyses showed that the variance explained at the teacher level was negligible (intra-class correlations ranged between 0.01 and 0.07). Therefore, our final models consisted of two levels (interactions within students). We estimated 8 models (2 independent * 4 dependent variables). The results of the final, random intercept models are summarized in Table 36.5.

Looking first at the differences in need-supportive teaching towards students with, versus students without learning problems, the total score on need-supportive teaching was higher for students with learning problems ($t(1389) = 2.60, p < .01$). There was no difference in the level of autonomy support offered to students with, versus students without learning problems ($t(2058) = .44, p = .33$). The degree of structure offered by teachers was higher for students with learning problems ($t(1405) = 3.00, p < .01$). Similarly, we see a higher degree of involvement for students with learning problems ($t(2054) = 2.18, p < .01$).

Comparing students with behavioral problems to students without the problems, the pattern of results was somewhat comparable to the results for learning problems, but the observed effects were smaller and none were statistically significant. Although teachers also tended to provide a higher level of need-supportive teaching to students with behavioral problems, the difference is not significant ($t(1389) = 1.31, p = .10$). Again there was no difference in the level of autonomy support offered to students with, versus students without behavioral problems ($t(2058) = .45, p = .33$). The same holds true for the degree of structure offered in one-on-one interactions ($t(1405) = .93, p = .18$). Teachers tended to show a higher level of involvement towards students with behavioral problems compared to students without behavioral problems, but this trend was not significant ($t(2054) = 1.19, p = .12$).

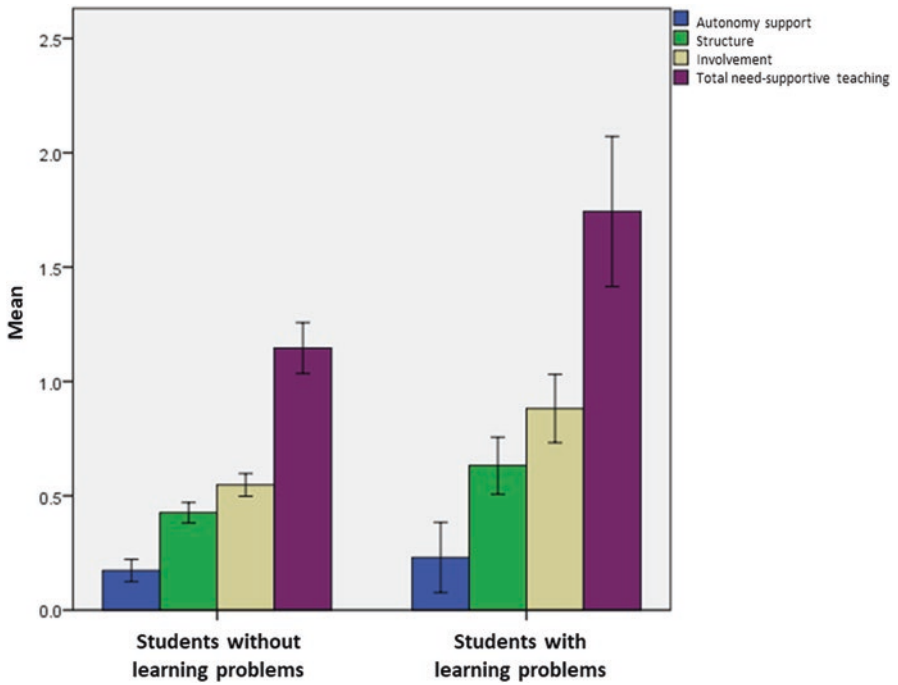


Fig. 36.1 Levels of autonomy support, structure, involvement and total need-supportive teaching towards students with and without learning problems

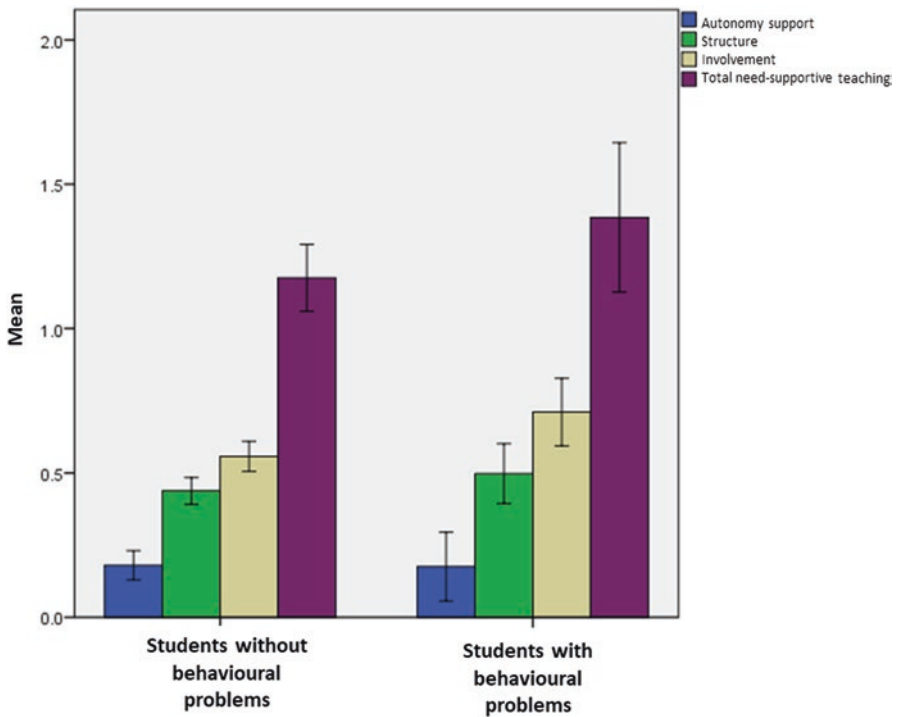


Fig. 36.2 Levels of autonomy support, structure, involvement and total need-supportive teaching towards students with and without behavioural problems

Table 36.5 Results of the multilevel models relating SEN (learning (LP) or behavioral problems (BP)) to dimensions of need-supportive teaching

	Dimension = Total need-supportive teaching			Dimension. = Autonomy support			Dimension. = Structure			Dimension. = Involvement		
	empty	LP	BP	empty	LP	BP	empty	LP	BP	empty	LP	BP
ICC	.10	.09	.10	.04	.04	.04	.03	.03	.03	.10	.09	.10
Coefficient (SE)		.64 (.25)	.27 (.20)		.04 (.09)	.03 (.07)		.24 (.08)	.06 (.07)		.24 (.11)	.11 (.09)
T-value (df)		2.60 (1389)	1.31 (1389)		.44 (2058)	.45 (2058)		3.00 (1405)	.93 (1405)		2.18 (2054)	1.19 (2054)
p-value		<.01	.10		.33	.33		<.01	.18		.01	.12

6 Discussion

From our data, we see small overall differences between the one-on-one interactions of teachers with students with and without SEN. Especially for students with learning problems, we see that teachers tend to show more involvement and an overall higher degree of need-supportive teaching. A similar (non-significant) trend is visible when comparing students with, versus students without behavioral problems. This does not align with previous research on the more often problematic teacher-student relationship when students have SEN, Based upon previous research on the more often problematic teacher-student relationship when students have SEN, one would expect a lower degree of need-supportive teaching. Next to the relatively small sample of teachers, perhaps this could have something to do with the fact that we used observations of interactions as they occurred at the very start of the school year, instead of the more aggregated impressions of closeness and conflict that teachers reported in questionnaires in previous studies (Zee et al., 2017). Teachers also provide more structure in interactions with students with learning problems, compared to students without learning problems. Offering structure in interactions with individual students means monitoring what students understand and adjusting instruction and feedback accordingly, which is what we also measured in our data. This kind of adaptive teaching is also a core element of differentiation (Deunk et al., 2018). The fact that the teachers in our sample did this, and to a larger extent for students who are known to have learning difficulties, is a positive indicator for their ability to differentiate instruction on a micro-level.

7 General Discussion: Linking Intentions to Differentiate to One-on-One Interactions

The aim of our two studies was to analyze teachers' intentions regarding differentiation on the one hand, while on the other hand examining the differences between one-on-one interactions with students with and without special educational needs.

In our two studies, we see on the one hand that teachers' often *do not formulate intentions* relating to differentiation between students with different educational needs or abilities. On the other hand, we see in the naturally occurring one-on-one interactions that teachers *do act* differently towards individual students with and without SEN, although these differences are small. Together, these two studies highlight two important aspects of teaching in general and differentiation in particular: pro-active planning of lessons on the one hand, and on the other hand the more improvisational skill of adjusting one's behavior and instruction from moment to moment in response to the emerging behavior of different students in the classroom (Sawyer, 2011). Differentiation is a particularly complex skill that can take a long time to master. Therefore, pro-active planning is considered a key element of differentiation (Tomlinson & Imbeau, 2010; Van Geel et al., 2019). It is in that sense worrying that only a small portion of teachers' intentions related to differentiation and that the intentions that did, were mostly formulated briefly and in very general terms. This might be an impediment towards actually implementing differentiation in the classroom.

Concerning teachers' actual behavior in one-on-one interactions, we see, however, that teachers in general show at least moderately positive levels of need-supportive teaching, and somewhat more towards students with SEN on some dimensions. As adaptive teaching is an important element of both need-supportive teaching as well as differentiation, this can be seen as a positive indicator of teachers' ability to differentiate in the 'improvisational' sense. However, we must emphasize that offering need support in individual interactions is, although a key condition, only part of differentiation practices in the classroom. We did not assess, for instance, whether teachers differentiate in the sense of grouping students according to ability, offering extra instruction time or adjusted goals for students with varying levels and needs or provide different assignments for different students. Two important goals for future research are therefore, first, to assess differentiation on the level of the whole lesson. Second, we studied intentions and teacher behavior currently in two separate studies. A logical next step would be to see whether we can predict teachers' actual differentiation practices from their intentions: is formulating detailed plans for differentiation in one's next lesson(s) a necessary prerequisite for implementing differentiation?

The added value of the studies presented here is that they inform us about the intra-individual level of differentiation. Although we investigated differentiation only in a relatively small sample of teachers, the intensive data collected provide a unique and ecologically valid insight into teachers' intentions as well as their behavior in interactions with students. This will allow us to make more detailed predictions of lesson-to-lesson differentiation in the future. Next to looking at differences between teachers in their teaching practices, we need to know more about why differentiation 'works' in some lessons and moments, but not in others. This will allow us to not only understand differentiation better at a fundamental level, but also provide 'differentiated' support for teachers who wish to improve their teaching skills.

7.1 Implications for Research and Practice

Given that we only described teachers' intentions relating to differentiation, future research needs to focus on to what extent intentions for differentiation relate to actual differentiation practices, both at the classroom as well as on the individual level. In teacher education and professionalization programs, more attention can be paid to teachers' intentions and lesson plans for differentiation. A third important implication of our study is that teachers can be made more aware of their intentions given the need for pro-active planning of differentiation practices.

Funding Acknowledgement This project was funded by NRO (the Netherlands Initiative for Educational Research), project no. 405-17-302.

References

- Blaauw, F. J., Van der Gaag, M. A., Snell, N. R., Emerencia, A. C., Kunnen, E. S., & De Jonge, P. (2019). The u-can-act platform: a tool to study intra-individual processes of early school leaving and its prevention using multiple informants. *Frontiers in Psychology, 10*, 1808.
- Burden, R. (2008). Is dyslexia necessarily associated with negative feelings of self-worth? A review and implications for future research. *Dyslexia, 14*(3), 188–196.
- De Boer, A., & Kuijper, S. (2021). Students' voices about the extra educational support they receive in regular education. *European Journal of Special Needs Education, 36*(4), 625–641.
- Deunk, M. I., Smale-Jacobse, A. E., de Boer, H., Doolaard, S., & Bosker, R. J. (2018). Effective differentiation practices: A systematic review and meta-analysis of studies on the cognitive effects of differentiation practices in primary education. *Educational Research Review, 24*, 31–54.
- Flick, U. (2009). Qualitative Methoden in der Evaluationsforschung. *Zeitschrift für qualitative Forschung, 10*(1), 9–18.
- Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development, 76*(5), 949–967.
- Keuning, T., & van Geel, M. (2021). Differentiated teaching with adaptive learning systems and teacher dashboards: The teacher still matters Most. *IEEE Transactions on Learning Technologies, 14*(2), 201–210.
- Ledoux, G., & Waslander, S. (2020). *Evaluatie Passend Onderwijs: eindrapport* [Evaluation Education that Fits: final report]. Kohnstamm Institute.
- Majorano, M., Brondino, M., Morelli, M., & Maes, M. (2017). Quality of relationship with parents and emotional autonomy as predictors of self concept and loneliness in adolescents with learning disabilities: The moderating role of the relationship with teachers. *Journal of Child and Family Studies, 26*(3), 690–700.
- Opoku, M. P., Cuskelly, M., Pedersen, S. J., & Rayner, C. S. (2020). Applying the theory of planned behaviour in assessments of teachers' intentions towards practicing inclusive education: A scoping review. *European Journal of Special Needs Education, 1*–16.
- Osch, L. V., Lechner, L., Reubsat, A., & Vries, H. D. (2010). From theory to practice: An explorative study into the instrumentality and specificity of implementation intentions. *Psychology and Health, 25*(3), 351–364.
- Rademaker, F., De Boer, A., Kupers, E., & Minnaert, A. (2020). Applying the contact theory in inclusive education: A systematic review on the impact of contact and information on the social participation of students with disabilities. *Frontiers in Education, 5*(21), 271.

- Roy, A., Guay, F., & Valois, P. (2013). Teaching to address diverse learning needs: Development and validation of a differentiated instruction scale. *International Journal of Inclusive Education*, 17(11), 1186–1204.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68.
- Savard, A., Joussemet, M., Pelletier, J. E., & Mageau, G. A. (2013). The benefits of autonomy support for adolescents with severe emotional and behavioral problems. *Motivation and Emotion*, 37(4), 688–700.
- Sawyer, R. K. (2011). What makes good teachers great? The artful balance of structure and improvisation. In *Structure and improvisation in creative teaching* (pp. 1–24). Cambridge University Press.
- Schwab, S. (2018). *Attitudes towards inclusive schooling: A study on students', teachers' and parents' attitudes*. Waxmann Verlag.
- Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4, 1–32.
- Smale-Jacobse, A. E., Meijer, A., Helms-Lorenz, M., & Maulana, R. (2019). Differentiated instruction in secondary education: A systematic review of research evidence. *Frontiers in Psychology*, 10, 2366.
- Stroet, K. (2014). Studying motivation in classrooms: effects of teaching practices on early adolescents' motivation.
- Stroet, K., Opendakker, M. C., & Minnaert, A. (2013). Effects of need supportive teaching on early adolescents' motivation and engagement: A review of the literature. *Educational Research Review*, 9, 65–87.
- Tomlinson, C. A., & Imbeau, M. B. (2010). *Leading and managing a differentiated classroom*. ASCD.
- Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., et al. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *Journal for the Education of the Gifted*, 27(2–3), 119–145.
- van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research*, 49(2), 127–152.
- van de Grift, W., Helms-Lorenz, M., & Maulana, R. (2014). Teaching skills of student teachers: Calibration of an evaluation instrument and its value in predicting student academic engagement. *Studies in Educational Evaluation*, 43, 150–159.
- van de Pol, J., Volman, M., Oort, F., & Beishuizen, J. (2015). The effects of scaffolding in the classroom: Support contingency and student independent working time in relation to student achievement, task effort and appreciation of support. *Instructional Science*, 43(5), 615–641.
- van Geel, M., Keuning, T., Frèrejean, J., Dolmans, D., van Merriënboer, J., & Visscher, A. J. (2019). Capturing the complexity of differentiated instruction. *School Effectiveness and School Improvement*, 30(1), 51–67.
- Van Mieghem, A., Verschueren, K., Petry, K., & Struyf, E. (2020). An analysis of research on inclusive education: A systematic search and meta review. *International Journal of Inclusive Education*, 24(6), 675–689.
- Yan, Z., & Sin, K. F. (2014). Inclusive education: Teachers' intentions and behaviour analysed from the viewpoint of the theory of planned behaviour. *International Journal of Inclusive Education*, 18(1), 72–85.
- Zee, M., de Jong, P. F., & Koomen, H. M. (2017). From externalizing student behavior to student-specific teacher self-efficacy: The role of teacher-perceived conflict and closeness in the student–teacher relationship. *Contemporary Educational Psychology*, 51, 37–50.

Elisa Kupers is associate professor at the Department of Inclusive and Special Needs Education. Her expertise is on teacher-student interactions in diverse and inclusive settings.

Anke de Boer is associate professor at the Department of Inclusive and Special Needs Education and Director of Educational Quality and Research at the RENN4 expertise centre for special education. Her expertise is on social inclusion, inclusive education and special education.

Judith Loopers is a PhD student at the University of Groningen on the project Differentiation Inside Out where she investigated differences students' daily experiences of motivational processes and teachers' differentiation skills.

Alianne Bakker is a PhD student at the University of Groningen on the project Differentiation Inside out where she investigated teachers' intentions and differentiation skills, and a teacher educator at the NHL Stenden University of Applied Sciences.

Alexander Minnaert is a full professor at the Department of Inclusive and Special Needs Education. His expertise is on learning and educational problems, school support and counseling, and inclusive education.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Epilogue

Klaas van Veen

This book reads like an international dialogue on teaching effectiveness, in which each chapter brings another perspective, definition or insights into the conversation. It also reads like a current overview and update of what is known from research in different settings and contexts. It confirms what was known, and at the same time new insights are developed. The studies from outside the Western countries are a very welcome addition to this dialogue. However, because of its focus on teaching in the classroom, two other issues are left out which are also crucial for teachers to be effective, namely (1) a deep understanding of how their students feel, develop and learn, and (2) how their work is organized and the time and space teachers have to really teach effectively.

In general, as stated in the introduction, is the focus on classroom processes or instructional practices related to student learning. More specifically on the classroom and what is happening there in terms of how teachers organize their teaching and all the factors that affect students' learning and their learning outcomes. Because of its focus on teaching effectiveness how students learn gets less attention. For teachers to be really effective in a classroom full of students, a deep understanding of how they learn is however essential. This refers not only to the development stage of students between 4 and 18 and their ability to learn, their social, practical, cognitive and metacognitive skills, but also to an understanding of how students learn, comprehend and gain specific domains of knowledges, skills, insights, and attitudes. What exactly is easy to learn and what is difficult for a 14 year old in a school setting? What are the common misconceptions, and what is their prior knowledge? And it refers also to an understanding of the world students live in. What inspires them? What challenges them? What is boring to them? Effective classroom management for example is largely a matter of organizing learning activities that are perceived by students as relevant and engaging, giving them a sense of structure and meaning (Doyle, 2006). To be effective in classroom management requires this deep knowledge of one's own students and how they learn and behave in a classroom. As research on expert teachers showed, expert teachers were especially effective in

class settings in which they knew their students very well, and not so much in the experimental settings in which they were asked to teach students unknown to them (Berliner).

Therefore, for teachers to learn to become effective, it is not only relevant to practice and experience all the different instructional practices and classroom processes, but also to gain this deep understanding of how students feel, develop and learn. Just like some of the observation instruments discussed in this book, most teacher education programs follow however Fuller's stages of concerns, in which beginning teachers are first focused on their selves, then on their tasks and finally on their impact on students (Fuller, 1969). To be able to perceive students' learning as a beginning teacher is hardly possible because of the survival modus they are in. Though Fuller never stated that this is the way teachers should develop, but how they develop once they learn to become teachers. To paraphrase Berliner, this might also be the way teachers developed if they are not deliberately educated. When teachers are purposely trained and guided, it would be good to start to educate them in understanding how their students feel, develop and learn. Followed by making them understand how you effectively can teach those students, so how to adapt your teaching to their learning and which classroom processes and instructional practices are functional in that context and situation.

As stated, the focus of teaching effectiveness is mainly on the classroom. Because of this focus on how teachers act in the classroom, the time and space they have to teach effectively is hardly explored. This refers to how the work of teachers is organized, both in terms of time and professional autonomy. To teach effectively, however, is related to the time and space one has. Moreover, the way the work of teaching is organized is largely based on how teaching is defined and perceived. If effective teaching is seen as adapting one's teaching to an understanding and insight in how your students feel, develop and learn, then seeing those students in classes of 30 students twice per week does not really enable teachers to teach effectively. If teaching often implicitly is seen as merely effectively organizing and teaching lessons, and effectively managing groups of students, then teaching classes of 30 students that you see twice a week is less of a problem to teach effectively.

In many countries the regular teaching is organized in relatively large groups (25–35), teachers teach 20–25 h per week, and there is relatively little structural time to analyze and to adapt one's teaching to their students' learning. Countries also differ in the manner of (collective) professional autonomy teachers formally have and experience in making decisions in their teaching. Training teachers in such working context to become more effective will be problematic because of the lack of time and space to develop and learn, actually a lack of time and space to teach effectively.

This time and space to teach effectively also seems a key in understanding the recurrent problems of a decrease in educational level and an increase in social inequality that many countries are dealing with. Many approaches can be found and among them is improving the quality of teachers, learning them to teach more effectively. Apparently, the teachers are largely to blame for these complex and strongly

social issues. It is still assumed that ‘schools can compensate for society’, to paraphrase Bernstein (1971)’s famous statement about the limited possibilities of schools and teachers to solve such issues. However, what possible solutions have in common to those problems is that those students are helped with individual attention on how they feel, develop and learn. That they are seen by teachers and that they feel they are seen. One very powerful factor in teaching effectiveness is having high expectations towards your students. Translating this concept into daily classroom practices implies that teachers have the time and space to see each student, to know them, to know what moves and drives that student, and to have a meaningful contact with those students. Otherwise those students won’t relate to the high expectations the teacher has of them.

Furthermore, the long-term effect of organizing teachers’ work in such a manner is that it strongly affects the social image of teaching. Teaching in this view is reduced to working with large groups, that you are supposed to manage and teach effectively in terms that disorder is avoided and the student outcomes are sufficient. Such a view makes teaching hardly attractive. It also largely explains the problems of increasing teacher shortages. For years, policy makers try to change this image of teaching by focusing on the joy of working with the younger generation, showing that effective teaching is an art or increasing teacher salaries. The problem of teacher shortages is still there and increased. The focus should be on the time and space to teach effectively, implying that teachers have sufficient time to analyze how their students feel, develop and learn, and based on those insights to adapt their teaching to their students.

References

- Berliner, D. C. (2004). Describing the behavior and documenting the accomplishments of expert teachers. *Bulletin of Science, Technology & Society*, 24(3), 200–212. <https://doi.org/10.1177/0270467604265535>.
- Bernstein, B (1971). Education cannot compensate for society. In B. R. Cosin, L. R. Dale, & G. M. Esland (Eds.), *School and society: A sociological reader* (pp. 64–69). Routledge and Kegan Paul.
- Doyle, W. (2006). Ecological approaches to classroom management. In C. Evertson & C. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 97–125). Erlbaum.
- Fuller, F. F. (1969). Concerns of teachers: A developmental conceptualization. *American Educational Research Journal*, 6, 207–226. <https://doi.org/10.3102/00028312006002207>

Concluding Thoughts

Rob Klassen, Ridwan Maulana , and Michelle Helms-Lorenz 

Teachers and the teaching they deliver play a pivotal role in the day-to-day life of children and adolescents, influencing fleeting states such as daily mood, but also longer-term social-emotional and academic growth. For many children and adolescents, the effectiveness of the teaching they encounter opens doors that would otherwise stay shut, and allows them the best chance to access the myriad opportunities that a high-quality education provides. Developing a better understanding of effective teaching in multiple contexts provides policy-makers, researchers, and practitioners with the tools to ensure that all children and adolescents will be offered more equitable opportunities to develop into healthy and productive members of society.

Over the course of this book we have attempted to pin down the moving target of effective teaching by looking at this complex and admittedly contested concept from a wide variety of perspectives from around the globe: from western, eastern, northern, southern contexts; from high-income countries and lower-income countries; from countries with rich histories of research on the concept to those with emerging research traditions. What constitutes ‘effective’ or ‘good’ or ‘high-impact’ teaching is not exactly the same from one country to another, or even one school to another, but what is agreed is that effective teaching invokes *change*. In some contexts, change is valued and defined as outcomes on test scores; in other contexts, change is viewed in terms of social-emotional growth in students’ lives; in other contexts, the change is seen primarily from a collective or community perspective. The universally shared perspective is that effective teaching results in change or growth, and that this growth is observable—and, for some, preferably measurable—and is directly and indirectly linked to the actions, attitudes, and actions of the

R. Klassen
Department of Education, University of York, York, UK

R. Maulana · M. Helms-Lorenz
Department of Teacher Education, University of Groningen, Groningen, The Netherlands

teacher. It is upon this premise—that teachers are crucial agents for change—that this book was written.

As noted in the Introduction, we believe that educational improvement requires an understanding of the systems, contexts, and individuals that shape effective teaching. At the beginning of the twentieth century, American researcher Pittenger (1917) recognized the multi-faceted influences on student learning, with teachers and effective teaching playing crucial roles, asserting that learning outcomes were “to no small a degree a joint product, due to influences flowing from all the teachers in the school, and from agencies outside the school” (p. 108). Most education scholars recognize that effective teaching is constructed through multiple influences—certainly not resting only on the shoulders of individual teachers—but through the interactions of cultural, political, and other social influences. Thus, it is crucial to our understanding of teaching effectiveness that we consider the concept from as many perspectives as possible, not relying on single cultural or national viewpoints.

We have seen in the last 100+ years that most of the scholarly contributions on effective teaching have come from Western settings, and especially from the United States, which has arguably been the world-leader in studying and disseminating effective teaching practices. We can confidently assert that this volume gives a voice to researchers often unheard from around the world to contribute to the discussion, and to test and develop new conceptualizations, frameworks, and instruments to measure teacher effectiveness. We can, for the first time in a single volume, explore insights on teacher effectiveness from five continents, giving us a much broader perspective on effectiveness than research from a single country. Many of the contributions involve cross-national collaborations that produce new insights: with chapters including collaborators working together from China and the UK (Chap. 7), the Netherlands and South Korea (Chap. 8), Germany and Hong Kong (Chap. 9), Hong Kong, the UK, China, and the Netherlands (Chap. 15), and a diverse web of co-authors working together on several chapters (i.e., Chaps. 17, 19, and 23) where five continents are represented. It is remarkable really, to bring together such a wide and representative community of researchers intent on improving education outcomes by building a better understanding of effective teaching.

The book includes many fine contributions that provide new perspectives on effective teaching, but the story is far from finished, and considerable areas of research remain under-developed. We propose four key questions that remain largely unanswered, and will benefit from the attention of new scholars setting out their programme of research.

First we ask, *What are the key outcomes delivered by effective teaching?* Pinpointing the outcomes of effective teaching will necessarily vary by context, and is inextricably linked to shared conceptualisations of the objectives of education systems. Defining effective teaching in relation to specific outcomes will help clarify the concept.

Second, we ask *Who benefits from effective teaching?* The simplistic answer, of course, is ‘everyone benefits’, but greater attention to understanding how teaching practices benefit particular groups of students is needed. In this book, we start by

exploring effective teaching in complex environments, with a particular focus in Sect. 5 on differentiation and adaptive teaching. However, more work is needed.

Third, we ask *How does effective teaching influence outcomes over time?* More work on understanding the teaching-outcomes relationship over time is needed, and we have only a few longitudinal studies on the topic.

Finally, we need to continue to ask, *What is the role of the teacher?* The present volume focuses on *effective teaching*, not *effective teachers*, and although vigorous debates continue about focusing on individual teacher characteristics, important questions remain about how individual teachers vary in important ways vis-à-vis delivering effective teaching. Our goal in this book has been to make a contribution to improving the quality and equity of education systems around the world by better understanding the role of teaching effectiveness from multiple perspectives. We have brought together a wide range of theoretical, empirical, methodological, and practical insights from a rich array of international settings. The authors contributing to the book bring sometimes contrasting theoretical and methodological approaches to answering key questions about effective teaching, but all share the goal of improving education systems and the learning experiences of children and adolescents around the world. We trust that the scholarly contributions of this volume will spur future researchers across the globe to consider devoting attention to the shared goal of building stronger education systems for the benefit of all.

Reference

- Pittenger, B. F. (1917). Problems of teacher measurement. *Journal of Educational Psychology*, 8, 103–110.