

The Enabling Power of Assessment 4

*Series Editor:* Claire Wyatt-Smith

Dany Laveault

Linda Allal *Editors*

---

# Assessment for Learning: Meeting the Challenge of Implementation

 Springer

# **The Enabling Power of Assessment**

Volume 4

## **Series editor**

Claire Wyatt-Smith, Faculty of Education and Arts, Australian Catholic University,  
Brisbane, Queensland, Australia

This series heralds the idea that new times call for new and different thinking about assessment and learning, the identities of teachers and students, and what is involved in using and creating new knowledge. Its scope is consistent with a view of assessment as inherently connected with cultural, social practices and contexts. Assessment is a shared enterprise where teachers and students come together to not only develop knowledge and skills, but also to use and create knowledge and identities. Working from this position, the series confronts some of the major educational assessment issues of our times.

More information about this series at <http://www.springer.com/series/13204>

Dany Laveault · Linda Allal  
Editors

# Assessment for Learning: Meeting the Challenge of Implementation

 Springer

*Editors*

Dany Laveault  
University of Ottawa  
Ottawa  
Canada

Linda Allal  
University of Geneva  
Geneva  
Switzerland

ISSN 2198-2643

The Enabling Power of Assessment

ISBN 978-3-319-39209-7

DOI 10.1007/978-3-319-39211-0

ISSN 2198-2651 (electronic)

ISBN 978-3-319-39211-0 (eBook)

Library of Congress Control Number: 2016943037

Springer Cham Heidelberg New York Dordrecht London

© Springer International Publishing Switzerland 2016

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

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, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature

The registered company is Springer International Publishing AG Switzerland

# Foreword

Let me tell you of my journey of discovery in the realm of classroom assessment and assessment for learning. As you will see, the teachers I found along the way and the lessons we learned together have led me directly to this moment and this book, edited by Dany Laveault and Linda Allal.

I was trained to an advanced level in modern and classical test theory and practice in my graduate program at Michigan State University (with Linda as my classmate incidentally). Fortunately, my doctoral studies were overseen by Robert L. Ebel, the one member of the faculty who held that teachers were the key assessors in American education. Ultimately, his influence would guide my career (more about that below). I moved on from there to teach principles of sound assessment practice at the University of Minnesota before going on to five years of complex application of psychometric theory as Director of test development at ACT in Iowa City, USA, creating and equating college admission tests. I felt I was succeeding in bringing my traditional measurement expertise to bear on behalf of student well-being and the improvement of the American educational system. I was wrong, and what happened next literally jolted me into reality. I became a dad.

As our munchkin entered school, wife Nancy and I had a front row seat—with deep personal involvement—as the assessment processes in her schools and classrooms began to have their effect on her learning life. As you might expect, we saw the good, the bad, and the ugly unfolding before our very eyes. One did not need a background in modern and classical test theory to realize how many things could go wrong in the realm of classroom assessment. The challenges were not complex. The danger of poor test quality became immediately apparent. There was little or no assessment expertise at hand to evaluate quality. I had had a sense of this problem since graduate school but had no sense of its depth or importance until it touched our family directly. My career direction was immediately and profoundly changed forever more.

I was sure I could help and began figure out how. My search for like-minded colleagues focused on classroom assessment in the USA came up empty. At this time, virtually no one in our US measurement community showed any interest in

understanding the emotional dynamics of assessment. But I did find and internalize the writings of Terry Crooks in New Zealand. And I discovered a strong professional network of interest in Canada. It became clear that I would need to reach out far beyond my homeland if I was to acquire the needed expertise. I began attending conferences from Vancouver to Calgary and back to Vancouver.

Soon, Nancy and I left the world of psychometrics behind for the sake of improving classroom assessment. We created the Assessment Training Institute (ATI) in an empty bedroom of our home in 1992. We planned it as a professional development company intended to advance the assessment literacy of practicing teachers and school leaders. Our mission was to improve the quality of classroom assessments by translating complex validity and reliability concepts into commonsense ideas that we could share with teachers using everyday language.

However, as this work progressed and as our daughter ascended through elementary school, I became increasingly aware, again with the help of my teachers from other lands, of a far more serious issue in assessment in American schools. The assessment processes that formed the heart of our attempts to motivate student learning were causing at least as many students to give up in hopelessness as were inspired to strive for academic success. In fact, because the mission of our schools was to sort students based on achievement measured using classroom assessments, major segments of our student population were supposed to give up in hopelessness. This was the motivational intent of our system of assessment. Once again, I could find no one in the USA who seemed to care about the dynamics of the assessment experience from the student's point of view.

Fortunately, it was at this time that Anne Davies from British Columbia, Canada, and Ruth Sutton of England entered my life. Again, the international assessment community came through for me. They helped me understand the power of student involvement in the ongoing classroom assessment process as a way to develop in students a sense of control over their own academic well-being—to help them maintain the confidence that success is within reach if they keep striving. With the help of my international mentors, I expanded the mission of the Assessment Training Institute from merely improving classroom assessment quality to also help teachers to master strategies for student-involved assessment. Still, I could find little interest in the US measurement community and so we pioneered on at ATI pretty much on our own. Fortunately, however, we did discover increasing interest among teachers and school leaders. Our professional development agenda filled up fast; our books, training materials, videos, and events were received with enthusiasm and were being used.

A few years later, I learned of the work of a research team at King's College London who had completed a comprehensive research review on the impact of formative assessment on student achievement and learning. Paul Black and Dylan Wiliam entered my learning life. It was as though the stars were continuing to align as never before. Their review, published in 1998, was the foundation of the conception of 'Assessment for Learning' proposed by the UK Assessment Reform Group. At ATI, we had been teaching lessons in student-involved assessment for several years just because they made such complete sense. Now, here were the

researchers who could back it up with solid research evidence. I literally devoured the lessons they offered. Furthermore, even more exciting was my discovery at this time that Linda Allal, my brilliant graduate colleague, now at the University of Geneva in Switzerland, was following the same career path as me—the use of classroom assessment as a teaching and learning tool. My international learning network added a welcomed new teacher.

By this time, several leaders in the US measurement community had come onboard. Lori Shepard at the University of Colorado was exploring ways to improve preservice teacher preparation in classroom assessment. Linda Darling Hammond was developing standards of sound practice at Stanford, as was Tom Guskey at Kentucky. Jim Popham had shifted the direction of his work and his leading introductory assessment textbook clearly toward teachers and day-to-day classroom assessment. Clearly, classroom assessment was emerging.

We all seemed to be learning so much from each other about an exciting new vision of excellence in assessment—the use of day-to-day classroom assessment as tool for promoting student confidence and achievement. Ruth Sutton and I arrived at the conclusion that we needed to bring this growing international community together. The synergy, we believed, would be very powerful. So we recruited teams of like-minded teacher educators, researchers, and policy makers from Canada, Continental Europe, New Zealand, the UK, and the USA to come to Chester, England, in September 2001 for three days of mutual teaching and learning. We prepared within our teams by collecting information to share and reflecting on key discussion topics. Then, we came together for the first time, and it was as though a new intellectual family had been born. Kindred spirits focused on the research base, matters of professional development on classroom assessment, and how policy might guide sound practice. We had so much to learn and so much work to do.

Since then, our collective journey to understanding has been rich indeed. Over the past decade and a half, we have come together repeatedly to teach and learn from each other in Portland, Oregon, USA, in 2005; Dunedin, New Zealand, in 2009; Bergen and Solstrand, Norway, in 2011; and Fredericton, Canada, in 2014. As our gatherings have evolved, new regions and nations have joined us. Indeed, the next generation of researchers, teachers, and policy makers has been added to our teams. Starting with the New Zealand event, each meeting is accompanied by a conference for local educators, so we can teach and learn from them. We will come together next in Brisbane, Australia, in 2016, and anticipation runs very high.

Following the Fredericton meeting, Dany Laveault and Linda Allal took the editorial lead in preparing this volume which collects and reports on the lessons we have learned since the Chester meeting. We concluded from the very beginning of our journey that we needed to develop our understanding of keys to success in classroom assessment and assessment for learning through sharply focused research. We needed to create and bring into action high-quality professional development experiences for teachers, school leaders, and policy makers. And we needed to promote the kinds of assessment and educational policies that guide practices that we know will promote each student's academic and emotional well-being. As you read on in this volume, you will see that these same themes



provide the organizational structure Dany, Linda, and the authors have used to pool our collective wisdom in 2016. Oh my God, as you will see, we have learned so much!

And, I believe the presentations offered herein reveal that our impact has been profound. Even the last bastion of obsessive belief in accountability testing—the USA—is awash with federal, state, and local research, policy, and professional development in classroom assessment and assessment for learning. Our once small community is strong and growing. But, again as you will see, we have much work yet to do as we pursue a new vision of excellence in assessment. This volume describes that vision in three parts: Assessment Policy Enactment in Education Systems, Professional Development and Collaborative Learning about Assessment, and Assessment Culture and the Co-Regulation of Learning. It is worthy of note that each chapter in this book concludes with suggestions and recommendations of ways of meeting the challenges of implementation.

My thanks to all who have contributed to our collective learning and impact. Let the work continue. Carry on.

Portland, OR, USA  
February 2016

Rick Stiggins

# Acknowledgments

Some challenges are more easily met than others. The preparation of this book presented challenges of its own that were overcome thanks to the quality of the contributions of the authors and the timeliness of their responses to our requests. We wish to thank the 33 authors, from 13 countries, who agreed to share their experience and reflections regarding the implementation of Assessment for Learning (AfL), the challenges they have encountered, and the ways of meeting these challenges. The knowledge gained from implementation of AfL across a wide spectrum of conditions in different countries allows better understanding of why the transition from theory and policy to classroom practice has been met with varying degrees of success and how it can be improved. We particularly appreciate the efforts made by the authors to draw—from their extensive experience—suggestions and recommendations that can be useful for an international audience of policy makers, professional development providers, and practitioners wishing to implement assessment for learning. In addition, we wish to thank Dr. Rick Stiggins for agreeing to write the Foreword to this book. Rick's foreword shares his personal journey as a long-standing advocate of classroom assessment in support of student learning and as the initiator of an international network that held its first invited symposium in 2001.

This book is the fourth in 'The Enabling Power of Assessment' series that Springer has launched under the expert guidance of the Series Editor, Prof. Claire Wyatt-Smith. We thank her for her support in the preparation of this volume. We also would like to acknowledge the contribution of Dr. Anne Davies and Prof. Ann Sherman who, along with Dany Laveault, organized the international symposium on Assessment for Learning held in Fredericton, Canada, in April 2014. The idea of this book originated in discussions with the participants at this symposium.

We would particularly like to thank Annemarie Keur, Associate Editor at Springer, for her encouragement and her expertise in guiding the preparation and production of this book. Finally, we wish to express our appreciation to the Faculty of Education at the University of Ottawa for the financial assistance it provided and to Dr. Scott Uzelman for the excellent quality of his proofreading and copy editing.

Since the concept of Assessment for Learning was first proposed by the Assessment Reform Group in the UK in 1999, an expanding network of policy makers, professional development providers, and researchers—working in collaboration with school leaders and classroom teachers—has moved the field forward in a variety of directions, as demonstrated by the contributions to this book.

# Contents

<b>1</b>	<b>Implementing Assessment for Learning: Theoretical and Practical Issues</b> . . . . .	<b>1</b>
	Dany Laveault and Linda Allal	
<b>Part I Assessment Policy Enactment in Education Systems</b>		
<b>2</b>	<b>Assessment Policy Enactment in Education Systems: A Few Reasons to Be Optimistic</b> . . . . .	<b>21</b>
	Dany Laveault	
<b>3</b>	<b>Making Meaning of Assessment Policy in Australia Through Teacher Assessment Conversations</b> . . . . .	<b>35</b>
	Lenore Adie and Jill Willis	
<b>4</b>	<b>Effective Enactment of Assessment for Learning and Student Diversity in Australia</b> . . . . .	<b>55</b>
	J. Joy Cumming and Fabienne M. Van der Kleij	
<b>5</b>	<b>Formative Assessment Policy and Its Enactment in the Philippines</b> . . . . .	<b>75</b>
	Patrick Griffin, Louie Cagasan, Esther Care, Alvin Vista and Fe Nava	
<b>6</b>	<b>Communication and Collaboration: The Heart of Coherent Policy and Practice in New Zealand Assessment</b> . . . . .	<b>93</b>
	Jenny Poskitt	
<b>7</b>	<b>More Than Good Intentions: Policy and Assessment for Learning in Scotland</b> . . . . .	<b>111</b>
	Ernest Spencer and Louise Hayward	

**Part II Professional Development and Collaborative Learning About Assessment**

**8 Building Capacity: Professional Development and Collaborative Learning About Assessment . . . . . 131**  
Dany Laveault

**9 Implementing Assessment for Learning in Canada: The Challenge of Teacher Professional Development . . . . . 145**  
Christopher DeLuca, Adelina Valiquette and Don A. Klinger

**10 Teachers’ Professional Development in the Context of Collaborative Research: Toward Practices of Collaborative Assessment for Learning in the Classroom . . . . . 161**  
Lucie Mottier Lopez and Fernando Morales Villabona

**11 Cooperative Learning About Assessment for Learning . . . . . 181**  
Kari Smith

**12 Developing Assessment for Learning Practice in a School Cluster: Primary and Secondary Teachers Learning Together . . . . . 199**  
Sue Swaffield, Roszalina Rawi and Amanda O’Shea

**13 Implementing High Quality Assessment for Learning: Mapping as a Professional Development Tool for Understanding the *What to Learn, Why to Learn It, and How to Learn It*. . . . . 219**  
Maria Araceli Ruiz-Primo

**14 Assessment for Learning: A Framework for Educators’ Professional Growth and Evaluation Cycles . . . . . 237**  
Anne Davies, Sandra Herbst and Ann Sherman

**Part III Assessment Culture and the Co-Regulation of Learning**

**15 The Co-Regulation of Student Learning in an Assessment for Learning Culture . . . . . 259**  
Linda Allal

**16 Assessment Culture Versus Testing Culture: The Impact on Assessment for Learning . . . . . 275**  
Menucha Birenbaum

**17 The Role of Classroom Assessment in Supporting Self-Regulated Learning . . . . . 293**  
Heidi Andrade and Susan M. Brookhart

**18 Scaffolding Self-Regulated Learning Through Self-Assessment and Peer Assessment: Guidelines for Classroom Implementation . . . . . 311**  
Ernesto Panadero, Anders Jonsson and Jan-Willem Strijbos

**19 Assessment for Learning: Co-Regulation *in* and *as* Student–Teacher Interaction . . . . . 327**  
Margaret Heritage

**20 Supporting Students’ Learning: From Teacher Regulation to Co-Regulation. . . . . 345**  
Louise Bourgeois

# Editors and Contributors

## About the Editors

**Dany Laveault** is a Professor Emeritus (2015) at the Faculty of Education, University of Ottawa, Canada. He served the Faculty as Vice Dean for Research from 1998 to 2002. He has also been an invited professor at Université Libre de Bruxelles four times since 1991. In 1995, he was awarded the Benoît Poulin Prize for outstanding contributions in educational evaluation. From 1993 to 2001, he directed the international journal *Mesure et évaluation en éducation*. From 2004 to 2015, he acted as an expert consultant for the Educational Quality and Accountability Office of Ontario (EQAO) and more recently for the Ministry of Education of Quebec. His writings in French and English have been published in over fifty scholarly journals and books and have garnered him international recognition as an expert in educational measurement and evaluation. The Faculty of Education granted him the Award for Excellence in Research in 2012. In 2014, DeBoeck released the third edition of his book on test theories. His research has focused on how assessment practices may be improved to develop students' self-regulation skills, including student self-motivation. He has published several papers on the role of student self-assessment in learning and on teachers' professional judgment as it relates to the norms and standards of the teaching profession regarding classroom assessment [http://education.uottawa.ca/lafaculte/professeurs/?ref=detail\\_prof&id=100633](http://education.uottawa.ca/lafaculte/professeurs/?ref=detail_prof&id=100633).

**Linda Allal** is a Professor Emeritus in the Faculty of Psychology and Education Sciences at the University of Geneva, Switzerland. She obtained her Ph.D. in educational psychology from Michigan State University. In a career spanning 33 years as professor at the University of Geneva, she held positions as Chair of the Education Sciences department and as Associate Dean of the Faculty. She served as President of the Association pour le développement des méthodologies d'évaluation-Europe (Europe Association for the Development of Evaluation Methodologies) and as European editor of the journal of this association. After her initial work on generalizability theory and its applications in educational measurement, she oriented her research primarily in two directions: the role of

classroom assessment in support of student learning and the processes of regulation and co-regulation of learning in the areas of writing and mathematics in elementary classrooms. She was awarded an honorary doctorate by the University of Liège, Belgium, in 2013, in recognition of her research and publications on the relations between teaching, learning, and assessment in classroom settings. Her recent publications in French and English concern the co-regulation of student learning, the role of teachers' professional judgment in summative assessment, and the ways students engage in classroom activities in the areas of writing and problem solving <http://unige.ch/fapse/people/allal>.

## About the Contributors

**Lenore Adie** is a Senior Research Fellow with the Assessment, Evaluation and Student Learning Research Program in the Learning Sciences Institute Australia, at the Australian Catholic University. Her research focuses on assessment and moderation processes as these contribute to supporting teachers' pedagogical practices and student learning. She has a further interest in the enactment of assessment policy and the validity of assessment processes. Sociocultural theories of learning are utilized within her work to interpret this dynamic context. Her research has generated new knowledge in the field of assessment focusing on quality in assessment practices and processes, in particular within systems of standards-referenced assessment <http://lsia.acu.edu.au/>.

**Heidi Andrade** is an Associate Professor of Educational Psychology and Methodology, and the Associate Dean for Academic Affairs at the School of Education, University at Albany—State University of New York, USA. Her research and teaching focus on the relationships between learning and assessment, with emphasis on student self-assessment and self-regulated learning. She has written numerous articles, including an award-winning article on rubrics for *Educational Leadership* (1997). She has edited or coedited several books on classroom assessment, including the *SAGE Handbook of Research on Classroom Assessment* (2013) and *The Handbook of Formative Assessment* (2010), and has edited or coedited special issues of *Theory Into Practice* (2009) and *Applied Measurement in Education* (2013). She has enjoyed a long-term working relationship with arts educators in New York City, with whom she has developed and implemented formative assessments for the arts [http://www.albany.edu/educational\\_psychology/68388.php](http://www.albany.edu/educational_psychology/68388.php).

**Menucha Birenbaum** is a Professor Emeritus of educational assessment at Tel Aviv University, Israel. She earned her Ph.D. in educational psychology (quantitative and evaluative research methodologies) at the University of Illinois at Urbana-Champaign, USA, and her MA in education (research methodologies, measurement, and evaluation) at Tel Aviv University. Menucha's current research focuses on formal and informal assessment for learning (AfL) and school-related contextual factors that facilitate or hinder the successful implementation of AfL, on



learning environments that promote learners' agency, and on large-scale diagnostic assessment. Menucha is currently leading a school-based professional development project in AfL for science and math teams in 15 middle schools, targeting their development as professional learning communities.

**Louise Bourgeois** is a Ph.D. student in the Faculty of Education at the University of Ottawa and currently teaches in the Faculty of Education at Laurentian University in Sudbury, ON, Canada. Her research interests center on teachers' professional judgment, assessment for learning, and collaboration as a form of professional development. In her doctoral dissertation, she focuses on how collaboration can influence the quality of teachers' assessment decisions to support student learning. She has 25 years of experiential knowledge in education as a teacher, a principal, and an education officer at the Ministry of Education in Ontario where she was instrumental in the development and publication of many documents on assessment exemplars.

**Susan M. Brookhart Ph.D.** is an educational consultant and author. She is also an adjunct faculty member in the School of Education at Duquesne University, Pittsburg, USA. Dr. Brookhart was the 2007–2009 Editor of *Educational Measurement: Issues and Practice* and is currently an Associate Editor of *Applied Measurement in Education*. She serves on the editorial boards of several journals and on several national advisory panels, including the Technical Advisory Committee for the College Board's Advanced Placement Program. She has been named the 2014 Jason Millman Scholar by the Consortium for Research on Educational Assessment and Teaching Effectiveness (CREATE) and is the recipient of the 2015 Samuel J. Messick Memorial Lecture Award from ETS/TOEFL. Dr. Brookhart's research interests include the role of both formative and summative classroom assessments in student motivation and achievement, the connection between classroom assessment and large-scale assessment, and grading.

**Louie Cagasan** is an academic with the Assessment, Curriculum and Technology Research Centre, based in the University of the Philippines. He has the responsibility for psychometric work associated with a longitudinal research study designed to track students' learning progress across different learning environments and is invited as resource speaker on topics related to test development. Louie is the lead research officer in a study of formative assessment in the Philippines.

**Esther Care** is currently a Senior Fellow of the Brookings Institute in Washington, USA. She is a Fellow of the University of Melbourne Graduate School of Education, Australia. During her tenure at the University of Melbourne, she was the Deputy Director of the Assessment Research Centre and currently represents the University as the interim Codirector of the Assessment, Curriculum and Technology Research Centre at the University of the Philippines. Her doctoral work focused on measurement of vocational interests and aptitudes and has extended her interests in the area of educational assessment, assessment of early literacy, and collaborative problem solving.

**J. Joy Cumming** is a Research Director of Assessment, Evaluation and Student Learning in the Learning Sciences Institute Australia (LSIA) at Australian Catholic University (<http://lsia.acu.edu.au/people/professor-joy-cumming/>). Recent publications include *Valuing Students with Impairments: International Comparisons of Practice in Educational Accountability* (Springer 2012). A former secondary school English and mathematics teacher, her major research focus is educational assessment and accountability. Completion of formal law studies, including admission as a lawyer in the Supreme Court in Australia in 2009, has informed her assessment research focus on equity and students with disabilities. She is leading a major Australian Research Council Discovery Project on working with teachers to provide adjustments in summative assessments for students with disability in mainstream subjects.

**Anne Davies** works with systems, schools, and educators in Canada and around the world, focusing on system, school, and classroom alignment and learning through a variety of lenses, including engaging students in assessment, making sound professional judgments, evaluation, and reporting. She has taught at all levels. Anne's graduate work focused on literacy, learning difficulties, and assessment. She has authored more than 30 books and multimedia resources in the area of classroom assessment. Anne's current research focuses on how leaders use assessment in the service of learning [www.connect2learning.com](http://www.connect2learning.com).

**Christopher DeLuca** is an Associate Professor and Graduate Faculty member in Classroom Assessment and Curriculum Studies at the Faculty of Education, Queen's University, Canada. Previously, Dr. DeLuca was an Assistant Professor at the University of South Florida (Tampa, USA) and worked in the area of policy research in London, England. Dr. DeLuca's research examines the complex intersection of curriculum and pedagogy as operating within frameworks of educational assessment and accountability. In particular, Dr. DeLuca's research centers on how preservice and in-service teachers learn to engage the complexities of assessing student learning in relation to the evolving accountability culture in today's classrooms. Dr. DeLuca's work has been published in national and international journals with continuous funding by the Social Sciences and Humanities Research Council of Canada. Dr. DeLuca is currently the President of the *Canadian Educational Researchers' Association* and Coeditor of the *Canadian Journal of Education*.

**Patrick Griffin** held the Chair of Education (Assessment) at The University of Melbourne, Australia, for almost 20 years as the Founding Director of the Assessment Research Centre. His work focuses on item response modeling applications in interpretive frameworks for performance assessment, problem-solving behavior, and higher-order competency assessment and performance reporting. He is a measurement team leader for UNESCO and the World Bank in Africa and Southeast Asia. He has led several national and international studies of problem solving, literacy, and numeracy and was the Executive Director of the Assessment and Teaching of 21st Century Skills project.

**Louise Hayward** is a Professor of Educational Assessment and Innovation in the University of Glasgow, Scotland, UK. For more than 20 years, Louise has worked with researchers, policy makers, and practitioners to offer children and young people in Scotland better life chances by bringing research, policy, and practice in assessment into closer alignment. She has researched and published extensively in the fields of assessment, social justice, and transformational change. She was a member of the UK Assessment Reform Group and is currently a member of the International Symposium of Assessment for Learning Researchers. Louise has been involved in research projects for a range of national and international funders, including major UK research councils and the European Union. Her most recent publication, *The Sage International Handbook on Curriculum, Assessment and Pedagogy* (Wyse, Hayward, and Pandya 2016), reflects her growing interest in exploring the interrelationship of curriculum, assessment, and pedagogy to promote sustainable change.

**Sandra Herbst** is a noted system leader, author, speaker, coach, consultant, and educator with extensive experience in assessment, leadership, and adult learning. Her expertise, informed by wide-ranging practice and research, is enhanced by her compassion and humor. She works with teachers, school and system leaders, and trustees as they plan and implement organizational innovation and strategic direction. Sandra has worked in both elementary and secondary schools as a classroom and specialty teacher, school administrator, and program consultant. She is the former assistant superintendent of the second-largest school district in Manitoba and a past President of the Manitoba Association of School Superintendents and the Manitoba ASCD Affiliate [www.connect2learning.com](http://www.connect2learning.com).

**Margaret Heritage** joined WestEd, a research, development, and service agency based in San Francisco, USA, as Senior Scientist in October 2014, after 12 years as an Assistant Director at the National Center for Research on Evaluation, Standards, and Student Testing (CRESST), at the University of California Los Angeles. Heritage's current work focuses on formative assessment, particularly as instances of co-regulation of student learning. In addition, she is conducting research on language learning progressions as the foundation for formative assessment with English learner students. Her work is published in peer-reviewed journals, edited books, and practitioner journals. Her latest book, *Formative Assessment in Practice: A Process of Inquiry and Action*, and her coauthored book, *English Language Learners and the New Standards*, are published by Harvard Education Press.

**Anders Jonsson** is a Professor of Education at Kristianstad University, Sweden. His main research interest is in classroom assessment, both for summative and for formative purposes, but he has also been responsible for the development of the Swedish National Assessments in biology, chemistry, and physics for year 6 in compulsory school. Currently, Anders is working on a research project funded by the Swedish Research Council, investigating how teachers handle the needs of 'borderline students' who are at risk of failing certain subjects in school. He is also,

together with Ernesto Panadero, exploring the relationship between Assessment for Learning (AfL) and Self-Regulated Learning (SRL), as well as how the use of scoring rubrics can support both AfL and SRL.

**Don A. Klinger** is a Professor in assessment and evaluation and the Associate Dean of Graduate Studies and Research at the Faculty of Education, Queen's University, Kingston, Canada. Dr. Klinger's research explores both classroom assessment and the psychometric and policy issues of large-scale assessments, program evaluation, and measures of school effectiveness. Dr. Klinger is particularly interested in the methods we use to evaluate students and the subsequent decisions, practices, and policies that arise from these assessment practices. Through his ongoing funding and research, Dr. Klinger has built strong research collaborations and communication between the research community and practicing educators. Dr. Klinger was the Cochair of the task force that published the American National Standards Institute–approved *Classroom Assessment Standards* for the Joint Committee on Standards for Educational Evaluation (JCSEE).

**Fernando Morales Villabona** is a Ph.D. student in the Faculty of Psychology and Education Sciences at the University of Geneva, Switzerland. Currently working as an Assistant of Professor Lucie Mottier Lopez, he is a member of the research group EReD (Evaluation, Régulation et Différenciation des apprentissages). His research interests center on assessment for learning, social interaction among students, and regulation of learning in educational contexts. In his doctoral dissertation, he focuses on practices of collaborative assessment for learning in the classroom and the ways they can support the development of interpersonal regulation of learning processes involving primary school students.

**Lucie Mottier Lopez** is an Associate Professor in the Faculty of Psychology and Education Sciences at the University of Geneva, Switzerland. Her research concerns assessment and the regulation of learning in educational contexts. She served two terms (2008–2012) as president of an international association in the area of educational assessment and evaluation (Association pour le développement des méthodologies de l'évaluation-Europe). She is the author of numerous articles and chapters on assessment. Her latest book (2015) concerns formative and summative assessment of student learning in the classroom context. Her research group has developed collaborative research aimed at producing knowledge about teachers' assessment beliefs and practices and at enhancing teachers' professional development in the area of assessment [www.unige.ch/fapsc/ered/index.html](http://www.unige.ch/fapsc/ered/index.html).

**Fe Nava** is a specialist in Educational Research and Evaluation, with interests also in educational psychology, psychological statistics, and measurement, in the College of Education at the University of the Philippines. She has acted as the coordinator of the assessment research program at the Assessment, Curriculum and Technology Research Centre at the University of the Philippines.

**Amanda O’Shea** is a Senior Lecturer in mathematics education at the University of Northampton, UK, where she leads the postgraduate certificate program for specialist teachers of primary mathematics. Amanda also teaches in the masters’ programs in mathematics education. Previously, she worked as a primary teacher for 12 years before moving on to teach and research mathematics education and assessment for learning at the University of Northampton. Amanda’s doctoral research focused on teachers’ conceptions and practices in assessment for learning and children’s developing autonomy in primary mathematics. She is currently researching early career teachers’ expectations of mathematics teaching and effective Continuing Professional Development within mathematics and assessment for learning.

**Ernesto Panadero** is a Researcher in the Faculty of Psychology at Universidad Autónoma de Madrid, Spain (<http://www.ernestopanadero.es>). His research focus is twofold. The first is to study the relationship between self-regulated learning and the use of formative assessment (self-assessment and peer assessment). Secondly, he seeks to understand and promote self-regulated learning at the individual level and socially shared regulated learning at the collaborative learning level. He is one of the coordinators of the Special Interest Group ‘Assessment and Evaluation’ of the European Association for Research on Learning and Instruction and a member of the Continental Europe team in the Assessment for Learning International Network (AfLIN).

**Jenny Poskitt** is a Senior Lecturer in the Institute of Education, at Massey University in New Zealand. Jenny’s interest in assessment began in her early career as a primary teacher, teaching children across the junior and middle school areas, and deepened in more than two decades of research and postgraduate teaching. Jenny has led New Zealand-wide research projects in assessment (e.g., exemplar development and evaluation of Assess to Learn professional learning) and participated in an OECD project. She contributes to various government- and ministry-level advisory groups and publishes internationally on assessment.

**Rosalina Rawi** is currently a Ph.D. student in the Faculty of Education at Cambridge University, UK. Her areas of research include Assessment for Learning and Malay Language Education. In Singapore, Roszalina lectures at the National Institute of Education where she works with postgraduate diploma, degree, and diploma in education teachers. Her specialization is assessment and Malay Language teaching pedagogy. She is also a practicum supervisor working closely with school leaders to develop the teaching practices of beginning teachers. Previously, Roszalina worked at various secondary schools in Singapore for eight years with the Ministry of Education. Her current research is a landmark study of Assessment for Learning practices of teachers within the Singapore Malay Language classroom.

**Maria Araceli Ruiz-Primo** is an Associate Professor at the Graduate School of Education, Stanford University, USA. She investigates teachers' assessment practices and the assessment of student learning at both the large-scale and classroom levels. Her recent research work focuses on the development and evaluation of instructionally sensitive assessments and instruments for evaluating formative assessment practices in the classroom. Recently, she coedited a special issue on assessment in the *Journal of Research in Science Teaching*. The work reported in this book was conducted while she was Director of the Laboratory of Educational Assessment, Research, and Innovation at the School of Education and Human Development, University of Colorado Denver.

**Ann Sherman** is currently the Dean of Education at the University of New Brunswick, Canada. Ann started her teaching career as a high school science and math teacher and moved to elementary school teaching. She earned a graduate degree in Leadership and then a M.Ed. in Curriculum and Instruction from the University of New Brunswick. She ended her public school teaching career as a school administrator before completing a Ph.D. at the University of Nottingham. She has taught in university settings since 1996 and continues to teach and research early learning, formative assessment, inquiry-based science, and the connections between all three.

**Kari Smith Ph.D.** is currently the Head of the Norwegian National Research School in Teacher Education (NAFOL) at the Norwegian University of Science and Technology, in Trondheim, Norway. She has acted as the head of teacher education programs abroad as well as at the University of Bergen, Norway. Her main research interests are teacher education, professional development, mentoring novice teachers, and assessment for and of learning. Kari Smith is active in the European Association for Research in Learning and Instruction (EARLI); she has been coordinator of the EARLI Special Interest Group 'Assessment and Evaluation' and, more recently, coordinator of the Special Interest Group 'Teaching and Teacher Education.' She has published widely and has given invited talks in Australia, New Zealand, China, Dubai, Korea, Singapore, Africa, USA, South America, Europe, Israel, and in her own country, Norway.

**Ernest Spencer** an Honorary Senior Research Fellow in the School of Education, University of Glasgow, Scotland, UK, has promoted, conducted research on, and written about 'assessment as part of teaching' and assessment for learning since the 1970s. He led major assessment projects at the Scottish Council for Research in Education (1970s–80s) and assessment policy and teachers' continuing professional learning as the National Specialist for Assessment for Her Majesty's Inspectors of Education Scotland during the development (1980s–2000) of assessment for learning, national monitoring, and national qualifications. More recent work has included promotion and evaluation of assessment for learning with OECD, Scottish government, Education Scotland, and local authorities, with particular emphasis on the coherence of the interrelationship of curriculum pedagogy and assessment in policy and in practice.

**Jan-Willem Strijbos** is a Full Professor in the Department of Educational Sciences at the University of Groningen, the Netherlands (<http://www.rug.nl/staff/j.w.strijbos/>). He received his masters' degree from the Radboud University Nijmegen in 1999 and his Ph.D. degree (with honors) from the Open University of the Netherlands in 2004. From 2005 to 2010, he was a Postdoctoral Researcher and Assistant Professor in the Institute for Child and Education Studies at Leiden University. From 2011 to 2016, he was a Professor in the Department of Psychology at the Ludwig-Maximilians-Universität München, Munich, Germany. He is a coordinator of the Special Interest Group 'Assessment & Evaluation' of the European Association for Research on Learning and Instruction. His research focuses on the design, implementation, and effectiveness of interactive learning practices (collaborative learning, peer assessment and feedback, learning communities) in physical and virtual settings.

**Sue Swaffield** is a Senior Lecturer at the University of Cambridge, Faculty of Education, UK, and Adjunct Research Fellow at Griffith University, Australia. She researches and teaches in the fields of educational leadership and assessment, building on two decades of school teaching and advisory work. Sue is an Associate Editor for *Professional Development in Education* and on the Editorial Advisory Board of *Assessment in Education: Principles, Policy and Practice*. Her 2011 article, 'Getting to the heart of authentic assessment for learning,' has consistently been one of the journal's five most read articles. Her books include *Unlocking Assessment*, and she has written the assessment chapters for five editions of Andrew Pollard's classic text *Reflective Teaching in Schools*. Sue is a past president and honorary life member of the Association for Achievement and Improvement through Assessment (AAIA), a Syndic of Cambridge University Press, and a Fellow of Wolfson College, Cambridge <https://www.educ.cam.ac.uk/people/staff/swaffield/>.

**Fabienne M. Van der Kleij** is a Research Fellow in Assessment, Evaluation and Student Learning research of the Learning Sciences Institute Australia (LSIA) at Australian Catholic University (<http://lsia.acu.edu.au/people/dr-fabienne-van-der-kleij/>). Her research covers various topics in the area of formative assessment. Recent and ongoing research has focused on, for example, formative assessment for diverse students, including students with severe disability, formative assessment and feedback, curriculum in the Australian context, and assessment for learning in classroom practice. A specific focus is on the role of technology in providing effective feedback for students as well as capturing teacher–student formative interactions.

**Adelina Valiquette** is a Ph.D. candidate in Assessment and Evaluation at Queen's University Faculty of Education, Canada, under the supervision of Dr. Christopher DeLuca. Adelina earned a M.Ed. degree from Queen's in 2015 and a B.Ed. from the University of Ottawa in 2009. Adelina has taught secondary mathematics and science in Ontario, Portugal, and Korea. These teaching experiences led her to research the application of formative assessment practices within diverse classroom environments. While currently working toward a Ph.D., Adelina continues to build her own capacity in assessment as a part-time classroom teacher.

**Jill Willis** is a Senior Lecturer in the Faculty of Education at Queensland University of Technology, Brisbane, Australia. Her teaching and research interests focus on understanding how teachers and leaders create opportunities within assessment practices to enable learner agency. She studies student and teacher perspectives of assessment using visual and digital feedback loops to understand how personal reflexivity can inform system change. With Dr. Lenore Adie, she has researched with teachers in Queensland schools to understand how new national assessment standards are interpreted into classroom assessment practices [http://eprints.qut.edu.au/view/person/Willis,\\_Jill.html](http://eprints.qut.edu.au/view/person/Willis,_Jill.html).

**Alvin Vista** is a Research Fellow at the University of Melbourne, Australia. His doctoral research was based on data from the Assessment and Learning Partnerships (ALP), a large-scale research program funded by the Australian Research Council, to look at factors explaining the growth in student achievement. He has been extensively involved in the development of the problem-solving assessments in ALP and the data analysis for the project. In partnership with the Assessment, Curriculum and Technology Research Centre in the Philippines, he is involved in an ongoing project that surveys formative assessment practices in Philippine public schools to identify areas that are amenable to systemic change. His research interests include test development, research methods, and large-scale assessments.



# Chapter 1

## Implementing Assessment for Learning: Theoretical and Practical Issues

Dany Laveault and Linda Allal

**Abstract** This chapter provides a general introduction to the topic of this book: namely, the conceptualization of assessment for learning (AfL) and the challenges of its implementation. It addresses theoretical issues, including the definition of assessment for learning and its relations with other concepts—in particular, the formative and summative functions of assessment. It discusses the characteristics of student learning to be considered in designing AfL, as well as the external constraints and other practical considerations that influence the implementation of AfL. In conclusion, it presents the structure of the book in three parts dealing with three interrelated aspects of AfL implementation: policy, professional development, and classroom practice.

### 1.1 The Enabling Power of Assessment

Can educational assessment ensure reporting on student learning and also provide means for supporting and enhancing student learning? This essential question is not new. It was at the heart of the well-known handbook published by Bloom et al. (1971). Since then, many researchers, policy makers, school leaders, professional development providers, and classroom teachers have attempted to devise and implement diverse forms of assessment aimed at supporting student learning. These efforts have shown that assessment can support learning through a variety of means and in a variety of circumstances but that many challenges are also encountered on the path towards implementation.

---

D. Laveault (✉)  
University of Ottawa, Ottawa, Canada  
e-mail: dany.laveault@uottawa.ca

L. Allal  
University of Geneva, Geneva, Switzerland

This book is the fourth volume in the Springer series: *The Enabling Power of Assessment*. It shares the concern of the first book (Wyatt-Smith et al. 2014) regarding the design of assessment for quality learning, but it highlights more specifically the challenges of implementation and the possible ways of meeting these challenges at the level of policy, professional development, and classroom practice. The contributions to this volume concern student learning and assessment in K-12 education (kindergarten through secondary school), as well as professional learning about assessment by teachers and school leaders who work in K-12 settings.

‘Assessment for learning’ (AfL) is now recognized as one of the most powerful strategies for supporting student learning (Hattie 2012). Over the last 20 years, its theoretical foundation, drawing on cognitive science, social constructivist theories of learning, and models of motivation, has been strengthened. Extensive empirical research, including comparative studies with high effect sizes (William 2010), has been published to document the impact of a diversity of AfL strategies on student learning. This theoretical and empirical research base has increased the credibility of AfL for educational policy makers and has led a number of jurisdictions across the world to make it a mandatory part of their educational evaluation policy. The implementation of AfL as a mandatory assessment practice has, however, been met with varying degrees of success. New challenges have emerged in terms of policy implementation (OECD 2011), implications for teacher professional development and classroom practice, as well as coordination with other assessment policies such as accountability based on large-scale testing. In certain instances, this has resulted in distorted or superficial incarnations of AfL (Looney 2011; Marshall and Drummond 2006).

For now, it seems that educators and learners have not been able to reap all the potential benefits of AfL and that expectations have not been fully met. There exists nevertheless a rich portfolio of professional experiences, evaluation studies, and empirical research to reflect upon in order to identify the factors that are likely to promote successful implementation of AfL. Research on the process of educational change (Fullan 2009), international studies such as those sponsored by the OECD (2013), meta-analyses on different topics (e.g., feedback, cooperative learning, visible learning), as well as classroom-based research in various countries, have all led to significant progress in understanding the conditions for successful AfL implementation.

Several recent publications have addressed the implementation of assessment for learning. The articles in a special issue of the journal *Assessment in Education: Principles, Policy & Practice* (Hopfenbeck and Stobart 2015) describe large-scale implementations at the national or regional level in a number of countries, with particular emphasis on the effects of accountability measures and summative assessment on the way AfL and formative assessment are implemented. An article by Birenbaum et al. (2015) provides a concise overview of international trends in the implementation of AfL policy and practice. The chapters of the present book provide new insights on three key aspects of implementation:

- the formulation and communication of assessment policy and the factors that account for varying degrees of successful policy enactment;
- the diverse types of professional development activities that allow teachers and school leaders to better understand assessment for learning and more actively promote new assessment practices at the school and classroom levels;
- the dimensions of assessment culture and the processes of co-regulation that foster or inhibit teacher and student engagement in assessment for learning as enacted in classrooms and schools.

The contributors to this book draw on their experience as researchers, professional development providers, and/or policy advisors in 13 countries: Australia, Canada, England, Germany, Israel, New Zealand, Norway, Philippines, Scotland, Spain, Sweden, Switzerland, and the United States. They have all worked closely with teachers and school leaders and this allows them to present well grounded evidence and analyses showing why the transition from theory and policy to classroom practice has met with varying degrees of success and how it can be improved. Studies regarding AfL are spread across the world in a diversity of publications in several languages (scholarly journals, government publications, local jurisdiction evaluations, etc.) and are thus not easily accessible by policy makers, school leaders, teacher educators, and—ultimately—classroom teachers. It is the aim of this book to reflect on this body of international experience and bring together new theoretical and practical insights regarding the key conditions for implementing assessment for learning, the challenges that occur, and the ways of meeting them.

## **1.2 Assessment for Learning and Its Relations with Other Concepts**

The concept of ‘assessment for learning’ was introduced by the UK Assessment Reform Group (ARG) in 1999 to describe the orientation that should be given to assessment practices developed by teachers in the classroom in order to support student learning. The well-known and widely disseminated ARG leaflet published in 2002 gave the following definition:

Assessment for learning is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there.

This proposal encountered similar perspectives in other countries leading to the emergence of an international network of researchers which has held a series of conferences on assessment for learning and its relations with classroom practice, teacher education, and assessment policy. Rick Stiggins, who was director of the Assessment Training Institute in Portland, USA, initiated the first conference, organized in Chester, UK, in 2001. Other conferences followed: in Portland,

Oregon, USA, in 2005; in Dunedin, New Zealand, in 2009; in Bergen and Solstrand, Norway, in 2011; in Fredericton, Canada, in 2014. Participants in the New Zealand conference formulated an updated definition of AfL that highlights its dynamic nature as an integral component of teaching and learning:

Assessment for learning is part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information from dialogue, demonstration and observation in ways that enhance on-going learning. (Klenowski 2009, p. 264)

AfL adopts the central aim of formative assessment, as proposed by Bloom et al. (1971) in their mastery learning model, in that it seeks to ensure adaptation of teaching and learning activities in ways that will enable students to attain intended learning outcomes of schooling. It advocates, however, a vision of formative assessment as an interactive process that:

- is embedded in ongoing instructional activities,
- fosters active student involvement in assessment (through self-assessment, peer assessment, and activities that engage students in metacognitive reflection about their learning),
- concerns not only cognitive aspects of learning but also affective and social aspects (motivation, attitudes toward learning, cooperation in learning).

In the initial ARG (1999) publication, and in many publications since then, a distinction is made between assessment *for* learning and assessment *of* learning. This distinction is often linked to the functions of classroom assessment defined in the Bloom et al. (1971) handbook. Assessment for learning is equated with formative assessment (FA) that supports student learning, while assessment *of* learning is associated with summative assessment (SA) aimed at establishing students' grades in report cards and at other forms of reporting for accountability purposes.

The assessment *for* versus *of* learning distinction is not, however, very satisfactory. Assessment *for* refers to a purpose or function of assessment, whereas assessment *of* pertains to an object that is assessed. When assessment is carried out *for* learning, as a purpose, it must inevitably start with assessment *of* one or more aspects of student learning, as an object, whether it be the learner's progression in a series of tasks, the learning outcomes attained, or the strategies and attitudes shown during a learning activity.

A third distinction—assessment *as* learning—has been introduced more recently. Without going into a detailed discussion here, it can be noted that two quite different meanings have been given to this concept. Seen in a positive perspective (Earl 2003), assessment *as* learning 'extend(s) the role of formative assessment for learning by emphasizing the role of the student, not only as contributor to the assessment and learning process, but also as the critical connector, between them' (p. 25). Seen in a more negative light, as expressed by Torrance (2007), assessment *as* learning refers to assessment tasks that masquerade as learning or supplant genuine learning activities.

One of the challenges of implementing assessment for learning has to do with clarifying what this expression means. Its definition may vary depending on the

assessment terminology adopted in a given country or educational jurisdiction. As will be seen in the chapters of this book, AfL has become a key concept of assessment policy in some countries whereas in other countries or jurisdictions the concept of formative assessment remains predominant. Other expressions are also used by a few authors: for example, ‘assessment for teaching’ referring to the use of assessment information by teachers to adapt instructional activities and provide feedback to learners (Chap. 5—Griffin et al.). There is nevertheless a great degree of overlap among the conceptions of assessment in support of student learning that are proposed and exemplified by the contributors to this book. This overlap is an underlying concern in our discussion, over the next three sections, regarding the conceptualization of assessment for learning.

### 1.3 Assessment for Learning: A Theoretical Definition Based on the Concept of Regulation

In order to account for how assessment information may be used to support and enhance learning in a variety of ways, we find the concept of ‘regulation of learning’ to be particularly insightful. This concept is not specific to the field of assessment; it is frequently used in a variety of disciplines (psychology, sociology, administration, computer science, to name a few). In psychology and education, the notion of regulation refers to the cognitive, social, and motivational mechanisms that govern changes occurring in learning and behavior. Well-known mechanisms postulated by learning theories include reinforcement in behaviorist theory, equilibration in Piaget’s constructivism, feedback loops in cognitive models, and social mediation in sociocultural and social constructivist approaches (Allal 2010). The role of regulation mechanisms has been explored in depth in contemporary socio-cognitive theory (Bandura 1997) and in the research on self-regulated learning (Zimmerman and Schunk 2011).

Allal (2010) has formulated the following generic definition of the regulation of learning:

Regulation involves four main processes: goal setting, monitoring progress toward the goal, interpretation of feedback derived from monitoring, and adjustment of goal-directed actions and/or of the definition of the goal itself. (p. 349)

This definition allows us to conceptualize assessment for learning as a family of assessment situations entailing processes of regulation which may be operationalized in a variety of ways. Each process may be made more or less explicit, may incorporate different sorts of tools, and may entail varying degrees of involvement on the part of the teacher and the students. In some cases, the teacher actively guides the processes of regulation, whereas in other cases student self-regulation or peer regulation are enhanced by the assessment procedures adopted.

The concept of regulation has long been a central feature of the French-language literature on formative assessment, with particular emphasis on the ‘interactive

regulation' of learning which results from students' interactions with teachers, peers, instructional materials, or assessment tools (Allal 1979). More recently, it has been argued that the regulations involved in classroom assessment can best be described as processes of 'co-regulation':

This means that student self-regulation develops in interaction with multiple sources of regulation in the learning environment and, at the same time, contributes to the deployment and exploitation of these sources in the learning activities undertaken in class. (Allal 2010, p. 349)

Hadwin and Oshige (2011) have analyzed how student self-regulation can develop in situations of co-regulation (which entail scaffolding by a teacher or more advanced peer) and in situations of socially shared regulation (where regulation is implemented collaboratively by the members of a peer group).

While different incarnations of AfL may be possible depending on the interactions between teachers and students, and among students, some forms of AfL may not flourish or even be possible depending on the level of control which teachers and students are able to exert on the processes of regulation in their educational environment. The implementation of different forms of AfL is influenced by educational policies and assessment frameworks which shape the context for teachers' professional development and their collaborative learning about AfL. Different combinations of external factors, with a variety of emphases on the processes of regulation, may lead to more or less successful, adaptive occurrences of AfL. For instance, what good is it to mandate AfL as part of an official assessment framework and policy if there is no appropriate professional development to help teachers acquire expertise in AfL? Similarly, how can teacher professional development foster AfL if student self-assessment is not included in the official assessment framework?

Because of the large number of possible AfL incarnations, implementation challenges are consequently diverse. Many of them may be subsumed under the label of AfL oversimplification. This happens each time one or more of the four essential processes of regulation is neglected or misconstrued, resulting in *mis-regulation* (Baumeister et al. 1994).

One regularly encountered instance of AfL oversimplification occurs with feedback. The way feedback is provided or used is not always effective. Ramaprasad (1983) defined feedback as '... information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way' (p. 4). Sadler (1989) emphasized that learners must apprehend both the 'reference level'—namely, the goal of their learning—and the actual level of their understanding. He pointed out that 'if the information is simply recorded, passed to a third party who lacks either the knowledge or the power to change the outcome, or is too deeply coded (for example, as a summary grade given by the teacher) to lead to appropriate action, the control loop cannot be closed and "dangling data" substituted for effective feedback' (p. 121).

The case of feedback is just one example of how important it is for AfL to attend simultaneously to all four processes of the regulation of learning. The coordination of these processes is essential to support learning effectively.

## 1.4 Assessment for Learning: A Practical and Operational Definition

In behavioral theories of learning, positive reinforcement is defined as a stimulus which will increase the odds that a specific behavior will be produced and maintained. Similarly, AfL needs to be held to some operational standard of outcome. While AfL cannot guarantee that learning will occur, it should at least increase its probability of occurrence. Hence the following definition:

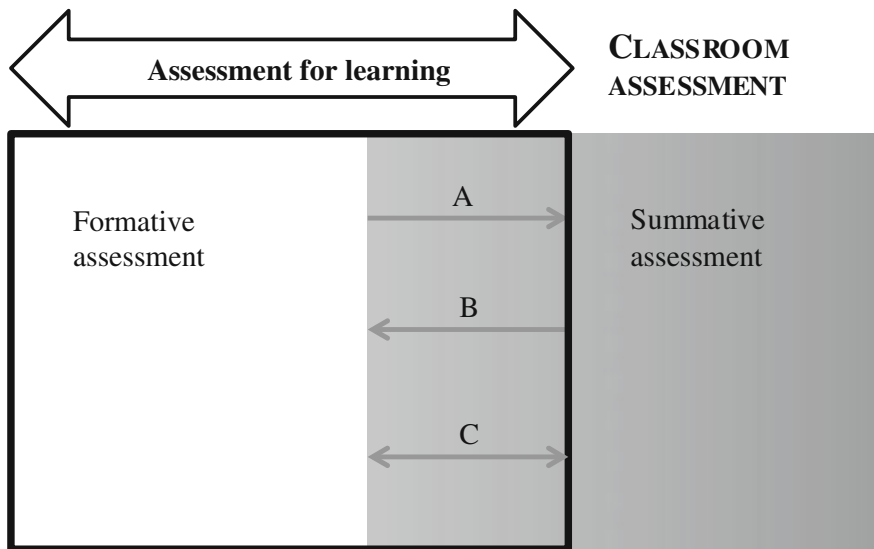
Assessment for Learning is the collection and interpretation of assessment information whose intentional use enables teachers and students, acting individually or interactively, to reach decisions that have a positive impact on instruction and learning.

Basically, as was the case for the definition of reinforcement, the above definition stresses the issue of effectiveness, whether assessment makes a difference in a student's learning progression or not. For example, when the use of assessment information leads to the choice of an instructional strategy that is more effective than an alternative one, this use of assessment information may be appropriately called AfL, independently of the form of the information or of how it was obtained—through anecdotal evidence, classroom observation, or formal exams—as long as it fosters student learning.

In this perspective, nothing prevents some forms of what is traditionally termed 'summative assessment' to meet such a criterion. According to Good (2011), there is value in both formative assessment and summative assessment: 'the challenge ahead of us is to put into practice the presumption that the label applied to an assessment is far less important than what is done with the information gathered' (p. 4). Indeed, information gathered on student learning may serve several different purposes, whether it is to *report on* student learning or *support* student learning, or both purposes at the same time. Consequently, Good (2011) proposes to use the expression 'formative use of assessment information' (p. 1) to emphasize the idea that the utilization of assessment information determines its function.

## 1.5 Assessment Synergies in Support of Student Learning

The idea of possible synergies between assessment for learning and some aspects of summative assessment, as carried out by teachers in their classrooms, has been addressed in a number of publications (Black et al. 2003; Harlen 2005; Stiggins 2008; Wiliam 2000). Observation of classroom practice and discussions with teachers show that there is often a zone of overlap between the formative assessments they practice and seek to promote, and the summative assessments they are required to carry out by the school system. The conceptualization of AfL therefore needs to take into account the reasons that lead experienced teachers to build



- A: information from formative assessment is taken into account in determining a summative assessment
- B: information from a summative assessment is used in a formative manner to support learning
- C: an assessment activity is composed of phases or components some of which have a formative function and others a summative function

**Fig. 1.1** Assessment for learning: its formative function and its interface with summative assessment (Allal 2011)

bridges between formative and summative assessment and the ways in which they build these bridges.

Figure 1.1 depicts the functions of ‘classroom assessment’ as developed and implemented by teachers (Allal 2011). The formative and summative functions appear as two overlapping zones. ‘Assessment for learning’ is shown as encompassing formative assessment but also several aspects of the interface between formative and summative assessment.

This interface includes three key practices (indicated by the arrows A, B, C in Fig. 1.1) that have been observed in research and in professional development activities with teachers.

- (A) *Information from formative assessment is taken into account in determining a summative assessment.* For example, when deciding on the mathematics grade to indicate in the student’s report card, the teacher takes into account summative test results but also observations and discussions with the student during formative assessments integrated in mathematics workshops during the semester. This use of formative information is seen as making the final



summative assessment more robust and valid than would be the case if the teacher simply calculated the average of the mathematics test scores (Allal 2013).

- (B) *Information from a summative assessment is used in a formative manner to support further learning.* For example, after grading students' biology reports (which will be included in the end-of-term grade in the students' report cards), the teacher analyzes the errors that frequently occurred in order to design new activities that will help the students progress in their understanding of key scientific concepts that were studied. This formative use of summative data is particularly prevalent when the number of teaching hours per week, with any given class, is quite limited and leaves little leeway for separation of formative and summative assessments. Black et al. (2003) found this practice to be one of the key ways in which teachers attempt to support the progression of student learning.
- (C) *An assessment activity includes phases or components with a formative function in support of learning, as well as phases or components with a summative function.* This refers to complex assessment activities that actively foster student learning and at the same time contribute to reporting on student learning. An example would be the preparation, over the course of a semester, of a writing portfolio that will determine the student's final grade but which, in each succeeding phase, engages the student in improvement of writing (through teacher and peer feedback) and in reflection on the processes of writing and revision (through metacognitive self-assessment). The portfolio procedures developed by Tierney et al. (1991) illustrate the linkages between formative and summative aspects of assessment in the areas of reading and writing.

We consider that the *optimal* forms of assessment for learning are those that are embedded in instructional activities and are designed to ensure interactive regulation of student learning, in coherence with the definition of the New Zealand AfL conference cited in Sect. 1.2. While seeking to promote these forms of assessment that have a clearly formative function (corresponding to the zone on the left in Fig. 1.1), we think it is nevertheless important to recognize practices in the interface with summative assessment which reflect teachers' efforts to coordinate their commitment to student learning with their institutional obligations in terms of grading and reporting.

There remain, however, many forms of summative assessment that may provide useful reporting information but do not support student learning (summative zone on the right in Fig. 1.1). For example, norm-referenced summative assessments, which involve the ranking of students according to their achievement, often have punitive outcomes (for the lower-ranked students) that inhibit rather than enhance classroom learning. Even very ordinary summative practices (e.g., grades communicated in students' report cards) can have negative effects, such as the distribution of rewards and punishments by some parents. These practices can affect the

goal orientation of students (Dweck 2000) and therefore their way of engaging in learning.

For summative assessment to support learning, careful and purposeful assessment design is required to collect and aggregate assessment information in such a way it will make up a *sum* that is more meaningful than its parts. In the absence of a design for combining assessments of student skills into meaningful categories, it is unlikely that classroom assessment—whether its initial purpose was summative or formative—can support student learning. This is what Shepard (2005) referred to as the ‘1000 mini-lessons’ problem where teachers try to address learning difficulties at the item level without putting such difficulties in a larger context. The appropriate use of assessment information raises important questions as to how assessment information is obtained, through appropriate and purposeful task design, and how it is structured, organized, and combined in a meaningful and useful way. According to Good (2011), ‘labeling an assessment item or activity as summative or formative without considering the timing and use can be misleading regardless of the quality of the item or the connection to instruction’ (p. 2).

## 1.6 Assessment for What Type of Learning?

One difference between assessment *of* learning (or summative assessment) and assessment *for* learning (or formative assessment) has to do with the degree of ‘elaborateness’ of assessment categories (Taras 2005). Let us take, as an example, one of the learning expectations in the area of writing (grade 4) from the Language curriculum of Ontario (Ministry of Education of Ontario 2006):

**Organizing Ideas.** Identify and order main ideas and supporting details and group them into units that could be used to develop a summary, using a variety of graphic organizers (e.g., a Venn diagram, a paragraph frame) and organizational patterns (e.g., generalization with supporting information, cause and effect). (p. 86)

In this example, the content in parentheses may be considered as an ‘elaboration’ of the learning expectation ‘Organizing Ideas’ so that they can be used when writing a summary. Such elaborations are suggestions used to exemplify the intermediate steps in the sequence of what the student is expected to learn. In some instances, ‘micro-summative assessments’ of learning may be carried out at a high level of elaborateness as a way of monitoring students’ progression before moving on to a more complex learning target. At the end of a learning cycle or a learning sequence, however, summative tests need to aggregate learning achievements in order to report globally on a student’s learning outcomes with respect to higher order learning expectancies. Reporting at a high level of elaborateness would not provide the student and the teacher with the global picture they need to set up new learning priorities and achievement targets.

In practice, a balance must be found between the need to aggregate assessment information for the purpose of reporting on student learning and the need to differentiate assessment information to support learning. The elaborateness of assessment information categories may thus play a role in both assessment of learning and assessment for learning. It will indirectly determine the degree to which an assessment will be holistic or analytical, thus impacting the level of specificity and timeliness of feedback.

Wiliam and Thompson (2008) go as far as saying that ‘virtually any assessment can be formative, provided it is used to make instructional adjustments and that a crucial difference between different assessments is the length of the adjustment cycle’ (p. 71). Both elaborateness and the length of the adjustment cycle define significant levels of domain specification and precision that are central in the design and use of assessment information. They will be referred to here as *fine-grained* assessment (high level of elaborateness and short time cycle) and *large-grained* assessment (low level of elaborateness and long time cycle).

The *large-grained/fine-grained* distinction (Laveault 2013) raises the issue of what types of learning are concerned when implementing AfL. This is relevant not only for deciding what can be combined within or across domains to report on student learning but also for deciding what kinds of information should be collected or interpreted to improve decision making about a student’s progression. Traditionally, taxonomies of learning objectives and curriculum specifications of contents and strands have been used to structure assessment information in a way that it can be transmitted and communicated by report cards. Content level domain specifications—at whatever level of granularity—have had, however, limited value in AfL because, in many instances, they are not aligned to learning theories (James 2006) that would help determine the nature of the learning difficulties, as well as the next steps in a student’s learning progression.

In teachers’ efforts to assist student learning, they need to consider and gather information along three main dimensions of learning:

1. *Depth of learning*, which can be assessed by gathering information on activities at different levels of complexity;
2. *Autonomy of learning*, which requires that assessments be collected in contexts where there is more or less guidance or support of student learning;
3. *Transfer of learning*, which requires assessment information collected in a variety of contexts to determine if learning generalizes to different and/or novel situations.

Whether fine-grained or large-grained, AfL should consider the distal or proximal nature of instructional goals. Overall expectations such as ‘generate, gather, and organize ideas and information to write for an intended purpose and audience’ (Ministry of Education of Ontario 2006) are long-term or distal goals which can only be achieved once some proximal goals, which are more specific and intermediate in a learning progression, have been reached at a certain mastery level.

Here are a few examples of intermediate goals taken from the Ontario curriculum (Ministry of Education of Ontario 2006):

- Developing ideas. Generate ideas about a potential topic using a variety of strategies and resources (e.g., brainstorm; formulate and ask questions to identify personal experiences, prior knowledge, and information needs)
- Classifying Ideas. Sort and classify ideas and information for their writing in a variety of ways (e.g., by underlining key words and phrases; by using graphic and print organizers such as mind maps, concept maps, timelines, jot notes, bulleted lists) (p. 86)

In the context of a classroom conversation about proximal goals, feedback may often concern the task (corrective feedback) or the self (use of praise as social reinforcement), which are the least powerful forms of feedback (Hattie and Timperley 2007) and may therefore not lead to long-term changes. While immediate feedback is often preferable to delayed feedback, more time for reflection may be required in the case of major distal goals involving vertical transfers of learning. For instance, it is much easier to provide immediate feedback during a student–teacher conversation about use of a graphic organizer—such as a Venn diagram—than to provide feedback on whether a student has gathered ideas to meet the intended purpose and audience. Feedback on self-regulation or feedback on cognitive processes, which are considered to be the most powerful forms of feedback (Hattie and Timperley 2007), may not be possible in a short classroom conversation; the teacher may need to read the whole piece of student writing, outside the classroom, in order to formulate useful comments and feedback, either in writing or as a plan for a future conversation with the student.

Large-grained assessments—as typically occurring in summative assessment—tend unfortunately to be constrained to retrospective information (Smith and Smith 2014). This restriction, as well as the tendency to limit AfL to proximal goals and fine-grained assessment, may narrow teachers’ capacity to be flexible and creative in the development and use of a variety of sources of assessment information. Here are a few examples of how some of these restrictions could be lifted to allow AfL to happen in a larger number of learning situations (Laveault 2013, 2014):

- Large-grained assessments make sense to the extent that they properly add up student achievements in similar domains or under similar conditions. For instance, lowering a student’s mark because an assignment has been turned in late makes it impossible to use the mark as prospective information because of the different nature of the skills added together. Assessing learning skills such as the capacity of the student to manage deadlines separately from subject matter achievement allows teachers to use information prospectively to set well-defined goals and to choose appropriate instructional strategies (Laveault 2008).
- While proximal goals help students to achieve long-term goals, these latter goals help students make sense of short-term ones. Prospective information gathered in large-grained assessments may be used to support students’ motivation by assisting them in setting challenging and realistic proximal goals leading to

attainment of long-term goals and to the development of an appropriate goal orientation toward learning (Dweck 2000).

- A better coordination of assessment information of different types may make better use of teachers' professional judgment (Laveault 2008). Teachers are often apprehensive about testing students when the information they have already obtained in the context of informal assessments or formative assessments lets them anticipate that the summative tests results will lead to a judgment of failure. In such cases, summative tests do not bring any new information and mainly confirm what the teacher already knows too well from interactions with students. It would seem more appropriate to use gradual or adaptive forms of assessment of learning—with more or less assistance or facilitating conditions—in order to report not only on learning outcomes but also on the level of assistance a student needs to succeed. Instead of eliminating all possible forms of assistance from summative tests, it might be preferable to factor them in to assess the whole extent of the student's level of learning. Such progressive procedures of summative testing have been described and advocated by several authors (Baxter and Glaser 1998; Laveault 2013; Rey et al. 2003). They allow information from formative assessment to be taken into account in determining a summative assessment and, reciprocally, to use summative information in a formative manner to support learning (as suggested in Fig. 1.1).
- Tasks may also be designed to determine students' degree of confidence in their own learning achievements (Leclercq 1993). Collecting information on students' confidence levels may help determine the extent they are able to set realistic goals, monitor their progression, critically reflect on their performance and seek help when they need it.

Integrating a variety of information sources into more flexible, mixed assessment designs may contribute to the development of synergies that would take into account length of learning cycle, domain precision, type of feedback, and goal proximity in order to improve AfL implementation in the classroom. Such mixed designs could help teachers improve the validity of the decisions they make as well as their capacity to report on and support three main dimensions of learning: depth, autonomy, and transfer.

## 1.7 External Constraints on Assessment for Learning

Classroom assessment can take a large number of forms depending on the ways the processes of regulation—target setting, feedback, monitoring of progress, adjustments of actions—are implemented and the degree of agency of both teachers and students in this implementation. The large number of potential occurrences of classroom assessment is limited however by factors in the educational environment and by the types of interfaces occurring between the formative and summative functions of assessment. Among these various conditions, external constraints on

teachers' and students' agency in the assessment process have a major influence. Here are some examples of different degrees of teacher involvement in assessment processes which may make it more or less possible to implement high quality AfL.

- *Assessment design and data collection.* Teachers may develop their own assessment tools—alone or as a group—or use assessment tools developed externally (e.g., assessment tools made available as part of instructional material; commercial tests, whether they are labelled ‘formative’ or not; large-scale tests administered more or less frequently by a local jurisdiction, etc.). Teachers may be in charge of collecting the information on their own students or they may be assigned to other groups, primarily in the case of large-scale testing. Official assessment policies sometimes require teachers to use results of large-scale, standardized tests to support student learning. This may, however, not be possible if the feedback provided to students is too global or arrives too late to allow regulation of learning. Furthermore, as teachers do not necessarily have a deep understanding of how such tests are built, it may be difficult for them to determine the diagnostic and formative value of the results. AfL is more likely to work well with instruments developed by the classroom teacher, or by a group of teachers, who then collect information on their own students.
- *Marking and interpretation of data.* In some large-scale testing, marking is done externally, but interpretation is left to teachers or to a group of teachers (acting as a professional community of practice). If interpretation of data is done externally, it can be difficult to find appropriate ways to support student learning since all the cumulative knowledge gathered on each student during a school year cannot be taken into consideration. Evaluation criteria may be developed by the teacher in interaction with colleagues—as in the case of ‘social moderation’ (Wyatt-Smith et al. 2010; Maxwell 2001)—and in interaction with students, especially in the case of classroom assessments. Criteria may also come with the assessment tools provided with the teaching material. Whatever method is used to develop and use criteria, AfL cannot be effective if students do not have a clear understanding of the standards and the criteria that will be used, by the teacher, by peers or by themselves.
- *Decision making.* The assessment information gathered on students’ performance may be used to support directly their learning progression, or serve as an indicator of the curriculum alignment with the program of study and of whole-school effectiveness. The follow-up on students’ results may be the responsibility of single teachers working with one or more students, or of teachers and other professionals working collaboratively. Decisions may be targeted at the whole school or at an entire classroom—as in interim assessment (Perie et al. 2009)—or at individual students’ learning.

While assessment may take a variety of forms, not all forms of assessment correspond to the goals of AfL. Some forms of assessment are much better adapted to the aim of AfL or can be used in a much more relevant and powerful way than others. Some forms of assessment are simply outside the realm of AfL (i.e., they

belong to the purely summative area on the right in Fig. 1.1). Others are targeted at whole classroom or school improvement rather than at supporting individual students' learning.

It is legitimate for school systems to use assessment information for the purpose of reporting and accountability and for the purpose of school improvement. In some cases, the results of external assessments may allow teachers and school leaders to identify problems of instruction affecting an entire school or several classes, and thus make some 'macro-regulations' that improve the quality of instruction and assessment practices: for example, ensure better alignment of teacher-constructed tasks and tests with curriculum standards, which may have a positive effect on student learning. These adjustments generally affect the conditions of learning and assessment for future cohorts of students, but they do not allow regulation of the learning processes in which the students—who were tested—are or were engaged (months before the external test was administered). It is a core principle of AfL that the evidence collected and used to make adjustments should benefit the students who provided the evidence.

System-wide and school improvements are very important goals and need to be pursued vigorously. But AfL has its own goals. It is important to find ways of coordinating improvements at the system and school levels with assessment for learning as it takes place within classrooms, in the actions of teachers and students. It is necessary to distinguish these arenas of action in order to coordinate their implementation.

## 1.8 The Structure of the Book

Since assessment for learning may take a variety of forms and is shaped by educational environments, its successful implementation needs to meet specific challenges at different levels of the education system, each level being intertwined with the others. Teachers' capacity to implement AfL depends on their access to high quality professional development activities designed to foster collaborative learning in interaction with other professionals: colleagues, school leaders, teacher educators, and specialists or researchers in the area of assessment. Teachers' new knowledge and skills will flourish, however, only if they fall on fertile ground. The successful implementation of AfL thus depends on school and classroom assessment culture and on the ways in which teachers transform newly learned skills into appropriate strategies for the co-regulation of student learning. Finally, the implementation factors involved in the two previous levels are largely conditioned by the assessment policy adopted by the system and by the resources devoted to its enactment.

Each *Part of this book* will focus on one level, while still taking into account relations with the other levels. The book is divided into three parts which address three topics defined as follows:

- Part I: Assessment Policy Enactment in Education Systems
- Part II: Professional Development and Collaborative Learning about Assessment
- Part III: Assessment Culture and the Co-Regulation of Learning

Each part begins with an introductory chapter written by one of the book editors. This chapter presents the topic to be dealt with, the contributing chapters, and the main suggestions and recommendations formulated in these chapters regarding the implementation of assessment in support of student learning. The authors' contributions and proposals are based on their extensive experience as policy consultants, professional development providers, and researchers working in close cooperation with teachers and school leaders. Although it is never possible to automatically transpose the lessons learned in one context (country or educational jurisdictions) to another, we believe that the pooling of ideas offers a basis for pushing forward the frontiers of knowledge about assessment for learning and for moving forward to meet the challenges of implementation.

## References

- Allal, L. (1979). Stratégies d'évaluation formative: Conceptions psycho-pédagogiques et modalités d'application. In L. Allal, J. Cardinet & P. Perrenoud (Eds.), *L'évaluation formative dans un enseignement différencié* (7th Ed., pp. 130–156). Bern: Lang.
- Allal, L. (2010). Assessment and the regulation of learning. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (3rd ed., pp. 348–352). Oxford: Elsevier.
- Allal, L. (2011, June). *Assessment for learning culture in the classroom, and beyond*. Paper presented at the 4th Assessment for Learning conference, Bergen, Norway.
- Allal, L. (2013). Teachers' professional judgement in assessment: A cognitive act and a socially situated practice. *Assessment in Education: Principles, Policy & Practice*, 20, 20–34.
- Assessment Reform Group (ARG). (1999). *Assessment for learning: Beyond the black box*. Cambridge: University of Cambridge.
- Assessment Reform Group (ARG). (2002). *Assessment for learning: 10 principles* (leaflet printed by the ARG).
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.
- Baumeister, R. F., Heatherton, T. F., & Tice, D. M. (1994). *Losing control: How and why people fail at self-regulation*. San Diego: Academic Press.
- Baxter, G. P., & Glaser, R. (1998). Investigating the cognitive complexity of science assessments. *Educational Measurement: Issues and Practice*, 17(3), 37–45.
- Birenbaum, M., DeLuca, C., Earl, L., Heritage, M., Klenowski, V., Smith, K., et al. (2015). International trends in the implementation of assessment for learning: Implications for policy and practice. *Policy Futures in Education*, 13(1), 117–140.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead: Open University Press.
- Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.



- Dweck, S. (2000). *Self-theories: Their role in motivation, personality and development*. Philadelphia, PA: Psychology Press.
- Earl, L. (2003). *Assessment as Learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press.
- Fullan, M. (2009). Large-scale reform comes of age. *Journal of Educational Change*, 10, 101–113.
- Good, R. (2011). Formative use of assessment information: It's a process, so let's say what we mean. *Practical Assessment Research & Evaluation*, 16(3), 1–6.
- Hadwin, A. F., & Oshige, M. (2011). Self-regulation, coregulation, and socially shared regulation: Exploring perspectives of social in self-regulated learning theory. *Teachers College Record*, 113(2), 240–264.
- Harlen, W. (2005). Teachers' summative practices and assessment for learning: Tensions and synergies. *The Curriculum Journal*, 16(2), 207–223.
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. New York: Routledge.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Hopfenbeck, T. N., & Stobart, G. (Eds.). (2015). Assessment for learning: Lessons learned from large-scale evaluations of implementation [Special issue]. *Assessment in Education: Principles, Policy & Practice*, 22(1).
- James, M. (2006). Assessment, teaching and theories of learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 47–60). London: Sage.
- Klenowski, V. (2009). Assessment for learning revisited: An Asian-Pacific perspective. *Assessment in Education: Principles, Policy & Practice*, 16(3), 263–268.
- Laveault, D. (2008). Le jugement professionnel d'évaluation scolaire: Enjeux, tensions et synergies nouvelles. *Revue Suisse des Sciences de l'Éducation*, 30(3), 483–500.
- Laveault, D. (2013, August). *Summative assessment for learning: How it may impact task design*. Paper presented at the 15th Conference of the European Association for Research on Learning and Instruction, Munich, Germany.
- Laveault, D. (2014). The power of learning-centered task design: An exercise in the application of the variation principle. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 109–121). Dordrecht: Springer.
- Leclercq, D. (1993). Validity, reliability and acuity of self-assessment in educational testing. In D. Leclercq & J. Bruno (Eds.), *Item banking: Interactive testing and self-assessment* (pp. 113–131). Dordrecht: Springer.
- Looney, J. W. (2011). *Integrating formative and summative assessment*. (OECD Education Working Papers, No. 58). Paris: OECD.
- Marshall, B., & Drummond, M. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education*, 21(2), 133–149.
- Maxwell, G. (2001). *Moderation of assessments in vocational education and training*. Brisbane: Department of Employment and Training.
- Ministry of Education of Ontario. (2006). *Language. The Ontario curriculum, grades 1–8*. Toronto: author.
- Organisation for Economic Co-operation and Development (OECD). (2011). *Evaluation and assessment frameworks for improving school outcomes. Common policy challenges*. Paris: OECD. <http://www.oecd.org/education/school/46927511.pdf>. Accessed September 20, 2014.
- Organisation for Economic Co-operation and Development (OECD). (2013). *Synergies for better learning. An international perspective on evaluation and assessment*. Paris: OECD. [http://www.oecd.org/edu/school/Evaluation\\_and\\_Assessment\\_Synthesis\\_Report.pdf](http://www.oecd.org/edu/school/Evaluation_and_Assessment_Synthesis_Report.pdf). Accessed July 23, 2015.
- Perie, M., Marion, S., & Gong, B. (2009). Moving toward a comprehensive assessment system: A framework for considering interim assessment. *Educational Measurement: Issues and practice*, 28(3), 5–13.
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28, 4–13.

- Rey, B., Carette, V., Defrance, A., & Kahn, S. (2003). *Les compétences à l'école*. Bruxelles: De Boeck.
- Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144.
- Shepard, L. (2005, October). *Formative assessment: Caveat emptor*. Paper presented at the Educational Testing Service Invitational Conference, New York, USA.
- Smith, J. K., & Smith, L. F. (2014). Developing assessment tasks. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing Assessment for Quality Learning* (pp. 123–136). Dordrecht: Springer.
- Stiggins, R. (2008). *Assessment manifesto: A call for the development of balanced assessment systems*. Portland, OR: ETS Assessment Training Institute.
- Taras, M. (2005). Assessment—summative and formative—some theoretical reflections. *British Journal of Educational Studies*, 53(4), 466–478.
- Tierney, R. J., Carter, M. A., & Desai, L. E. (1991). *Portfolio assessment in the reading-writing classroom*. Norwood, MA: Christopher-Gordon.
- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, 14(3), 281–294.
- William, D. (2000, August). *Integrating formative and summative assessment*. Paper presented at the International Congress on Mathematics Education, Tokyo, Japan.
- William, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment*. New York: Routledge.
- William, D. & Thompson, M. (2008). Integrating assessment with learning: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53–82). New York: Erlbaum.
- Wyatt-Smith, C., Klenowski, V., & Colbert, P. (Eds.). (2014). *Designing assessment for quality learning*. Dordrecht: Springer.
- Wyatt-Smith, C. M., Klenowski, V., & Gunn, S. (2010). The centrality of teachers' judgment practice in assessment: A study of standards in moderation. *Assessment in Education: Principles, Policy and Practice*, 17(1), 59–75.
- Zimmerman, B., & Schunk, D. (Eds.). (2011). *Handbook of self-regulation of learning and performance*. New York: Routledge.

**Part I**  
**Assessment Policy Enactment**  
**in Education Systems**

## Chapter 2

# Assessment Policy Enactment in Education Systems: A Few Reasons to Be Optimistic

Dany Laveault

**Abstract** This chapter presents the topic addressed in Part I: *Assessment Policy Enactment in Education Systems*. It starts from the position that while there is a lot to be learned about policy enactment, there are several reasons to be optimistic. It ties together the common policy challenges and directions faced by different education systems around the world before introducing new perspectives on policy implementation. The concept of co-regulation is the explanatory framework used to describe possible variations in models of policy implementation and to account for the challenges met by both policy designers and enactors. Several recommendations resulting from the evidence presented in the chapters in Part I are submitted to move forward in policy implementation. Some of them are already showing promising results.

### 2.1 Policy Implementation: Still Much to Be Learned

The manner in which policy is enacted is essential to a successful AfL implementation. Trochim (2009) defined evaluation policies as ‘any rule or principle that a group or organization uses to guide its decision and actions when doing evaluation’ (p. 16). These principles or rules are often disseminated as ‘official texts articulating the intentions of central authorities to guide the actions of participants’ (Ben Jaafar and Anderson 2007, p. 208). However, the enactment of such intentions is far from assured.

As is the case for any study on major changes occurring in education, those involving AfL implementation need time to acquire a reliable knowledge base we can count on to develop future policies. Not long ago, regarding AfL, Black and Wiliam (2003) wrote, ‘If we had restricted ourselves to only those policy implications that followed logically and inevitably from the research evidence, we would have been able to say very little’ (p. 628).

---

D. Laveault (✉)  
University of Ottawa, Ottawa, Canada  
e-mail: dany.laveault@uottawa.ca

More recently, Christie and Fierro (2012) observed an imbalance between the actual research base on evaluation policies implementation and their relative importance in educational systems:

Despite the central influence that evaluation policies may have in how evaluations are designed and conducted, and despite the fact that these policies exist in large federal programs, little empirical work has been conducted to better understand how they are interpreted and implemented by the evaluators and practitioners whose work they are likely to effect (p. 65).

The importance of monitoring the transition from policy adoption to policy implementation has been stressed by several authors (Christie and Fierro 2012). Considering the variations reported in the conceptualization and the operationalization of AfL, one can expect great variability in policy implementation. Thus, it appears all the more important to be able to monitor AfL implementation to ensure that it follows ‘both the letter and the spirit of AfL’ (Earl and Timperley 2014, p. 325).

## 2.2 Assessment Policy Enactments: A Regulation Conceptualization

According to Weinbaum and Supovitz (2010), ‘we learned that complex programs go through a process of “mutual adaptation” in which both developers and implementers make adjustments to work more effectively (Berman and McLaughlin 1978)’ (p. 68). Christie and Fierro (2012) found that the intentions of an evaluation policy are susceptible to change depending on how the policy and its underlying values are interpreted by those in charge of implementing it. We submit that such mutual adaptations can be explained and accurately described by a series of co-regulations, in a manner similar to the co-regulations that occur between teachers and students. The concept of co-regulation contributes to raise relevant questions regarding the implementation and enactment of AfL policies such as the following:

- *Policy Goals.* What kind of goals? How are they determined and by whom?
- *Sources and quality of feedback on policy enactment.* Feedback from whom? Students, teachers, school principals, parents? What makes for good feedback? Feedback on what?
- *Actions and agents.* Who are the actors/implementers? Who is accountable? Who makes decision and how? What actions must be taken to achieve the policy goals? What are the enablers and barriers of actions? What are the stakes for the actors/implementers?

In the case of implementing AfL policies, it appears that two layers of co-regulation are involved in a process of mutual adaptation:

- Co-regulation between the developers and the school leadership in charge of implanting policies at the school level.
- Co-regulation between the school leadership and the teachers in charge of implementing and enacting policies at the classroom level.

### 2.3 Variations in Policy Implementation Models

The educational environment of different jurisdictions will shape the manner in which AfL policies will be enacted depending on the co-regulations at play between policy designers and implementers at the school and at the classroom level. There are numerous ways AfL policies can be enacted, such as the following:

- *Variations in the way policy goals are determined.* The manner in which policy goals are determined, including top-down, bottom-up, and reciprocal interactions, will most likely impact on the intelligibility of goal interpretation and on the nature of engagement towards such goals.
- *Variations regarding the manner in which assessment information is circulated among different levels of the education system.* The kind of information and how much may be centralized or decentralized, or made public to a lesser or greater extent, will have an impact not only on matters of transparency but also regarding the utility of feedback to regulate future actions.
- *Variations regarding the policy accountability orientation.* Accountability is an integral part of performance-based policy development and implementation. Accountability is also an important factor regarding how AfL will be implemented. Spencer (2004 in Ben Jaafar and Anderson 2007) described two accountability orientations proposed by Blackmore (1988):
  - Policy targeted at improving the management of the school system: *Economic-Bureaucratic Accountability* (EBA)
  - Policy targeted at improving students' learning: *Ethical-Professional Accountability* (EPA)

AfL implementation fits in both an EBA or EPA orientation. Within the EBA orientation, the most advocated incentive to implement AfL is to increase students' performance and achievement levels through enhanced efficiency in the use of human and material resources. Teachers are directly held responsible for students' achievement results and therefore, should use AfL to improve them. Hence, in such a context, 'The results are what matters, and the processes are validated only by performance' (Ben Jaafar and Anderson 2007, p. 211). However, within the EPA orientation, the means are emphasized and responsibility is primarily collective. Emphasis is put on teachers working together as a professional learning community and on students' improved learning skills and sustained achievement levels (Ben Jaafar and Anderson 2007).

Policy accountability orientation may be determinant in AfL enactment because it has an impact on the extent teachers are individually or collectively held accountable for students' results, whether the means used to improve students' performance are important or not, and finally, whether teachers will be considered as autonomous professionals or 'semi-professionals' whose work need to be structured and closely supervised (Hodson and Sullivan 2012).

Depending on the educational context and the policy accountability orientation, AfL implementation will be different, more or less challenging and more or less likely to succeed. The successful implementation of AfL will depend on the answer to two questions of major importance:

To help both program designers and school-level implementers avoid the sense of failure, can we predict what parts of a program will 'stick' and what will be changed? Or can we identify the points at which adaptation is likely to take place? (Weinbaum and Supovitz 2010, p. 68)

To help foresee what can eventually work, we need to focus on the regularities across the various implementation contexts that have led to favourable outcomes. Each experience in AfL implementation, regardless of whether it was successful, has the potential to help us understand best practices. Hence, much can be learned from the comparison of various cases of AfL implementation regarding what makes a successful implementation of AfL and what kind of adaptations are necessary.

## 2.4 Assessment Policy Challenges

'One of the most consistent findings from education research is variability in program implementation' (Weinbaum and Supovitz 2010, p. 68). Such variability originates not only from the large variety of program designs but also from the manner in which they are interpreted and implemented. In their extensive review of the implementation of a California State evaluation policy that required grantees to conduct scientifically based research (SBR), Christie and Fierro (2012) found 'few projects were able to implement SBR projects in a manner consistent with the evaluation policy' (p. 71). Research directors of successful projects 'remained flexible and adjusted study designs as needed to accommodate contextual conditions' (p. 71). Christie and Fierro (2012) remarked that the flexibility shown in implementing the policy generated an unexpected result:

in the process of translating SBR into action, many studies ultimately contributed valuable formative information to local projects—an unintended outcome associated with funding SBR.... Findings from our study indicate that policy makers would be well served by embracing both the learning and accountability functions of evaluation in their evaluation policies (p. 72).

Non-conformity to initial policy design as well as flexibility in policy implementation may not only have important payoffs but may also be the sole means by which policies may be successfully implemented. Policy designers may benefit

from factoring in a degree of ‘discretion’ in policy implementation. Halverson and Clifford (2006) have defined discretion as ‘the actor’s power to use judgment to determine a course of action within the perceived constraints of a situation’ (p. 606). By including an opportunity for a degree of discretion instead of avoiding it, policy development and implementation may move beyond ‘discussion of policy fidelity’:

Policies are designed to constrain practitioners’ behaviour to produce intended practices and outcomes.... However, policies also rely on practitioner discretion to adjust policy demands to local circumstances or to fill in gaps left unspecified by policy design (Halverson and Clifford 2006, p. 606).

Thus, the degree of discretion allowed in policy implementation is a crucial element of policy implementation, whatever the policy targets. Schools have different needs, and implementing AfL may mean that policy targets may need to be adapted to consider the existing conditions of the school environment, such as the capacity of teachers to work together on issues of assessment as well as the school assessment culture that is already in place.

A space for a form of co-regulation must be saved to allow for appropriate accommodations to occur between policy designers and policy implementers. Halverson and Clifford (2006) identify two forms of discretion: *managerial discretion* and *learning discretion*. Managerial discretion is an essential component of the leadership expertise required in policy implementation, whereas cognitive discretion refers to the capacity to learn when opportunities are provided by the policy design.

The co-regulation of policy implementation appears to flow both ways: ‘both designers and practitioners need opportunities to learn from each other about (a) how policies are intended to change practices and (b) how practices need to inform policy development’ (Halverson and Clifford 2006, p. 608). Such co-regulations between designers and practitioners are more likely to occur when certain forms of interactions are built in the process of policy development and implementation, which would disqualify uniquely ‘top-down’ or ‘bottom-up’ forms of implementation.

The degree and kinds of discretion allowed in policy enactment are important conditions of successful AfL implementation. The decision to implement AfL has frequently been advocated and motivated by stressing performance enhancements, both on the part of teachers and of students. Such entrenched beliefs in the power of AfL may have blinded policy designers from the dual nature of AfL challenges: implementing AfL not only regards improving students’ and teachers’ *performance*, it also regards improving students’ and teachers’ *learning*. Seifert and Hutchins (1992) in Halverson and Clifford (2006), suggested that ‘it is much more difficult to design for learning than for system performance’ (p. 97).

A basic reflex of policy design is to base policy implementation on highly specific targets. Depending on the nature of the policy targets as well as of the context and the type of governance, it may be more or less worthwhile to achieve high levels of fidelity in AfL implementation. According to Weinbaum and Supovitz (2010), to target greater specificity does not necessarily increase the likelihood of fidelity in



implementation: ‘either finding may be true depending on the classroom, school, or district. However, the focus on increasing specification may distract from more important variables’ (p. 69).

To summarize, it appears that policy designers should use *discretion with discretion* when planning ahead for policy implementation. To meet the challenges of AfL implementation, policy designers need to foresee and consider how variable conditions may affect the enactment of the policy itself. Planning ahead opportunities for the occurrence of co-regulations among different actors of the school system may increase the odds of a successful implementation. Such advance planning requires that variations occurring in the school system be known and based on accurate facts. This planning also requires a certain theoretical and professional knowledge base to work out possible solutions at the local level to ensure successful implementation.

## 2.5 Common Policy Challenges and Directions: Lessons Derived from OECD Studies

From the large variety of potential situations that occur at the local level, the successful design and implementation of AfL would appear a nearly daunting task if it could not rely on the capacity of the school system to learn and adapt locally. There are reasons to be optimistic, and the implementation process may be made more predictable according to Weinbaum and Supovitz (2006):

Although adjustments are likely to occur at multiple places and repeatedly over time, the implementation process has junctures that can be identified and defined in ways that may increase the predictability of how programs are likely to be used (p. 69).

An OECD study has identified certain of these ‘junctures’ that prevent AFL from fully playing its intended function in a large number of jurisdictions. One important policy challenge is ‘to find suitable strategies that can integrate classroom-based formative assessment within the broader assessment and evaluation framework’ (OECD 2011, p. 5).

One of the major strategies needed to achieve such integration consists in developing a ‘closer interface between formative assessment and summative assessment’ (OECD 2011, p. 5). As previously noted in Chap. 1, there are several means by which this can be accomplished. Central to this strategy is ensuring that the assessment covers ‘the full range of goals set out in standards and curriculum over time and in a variety of contexts’ (OECD 2011, p. 5). To achieve this requires that each source of student assessment information be used optimally and to its full extent. For instance, teachers are in a strong position to follow students’ learning progression and to assess reasoning and problem solving through performance-based assessment on a continuous basis. Complementarily, standardized assessments provide an opportunity to validate teachers’ classroom observations and to help them obtain a better sense of the extent that their students’ achievements are appropriately aligned with the school curriculum at important

transition points. Although large-scale external assessment can barely provide the fine-grained information regarding students' learning difficulties, certain test banks, such as the *asTTle* (Assessment Tools for Teaching and Learning) and the Progressive Achievement Tests (PAT), which are used in New Zealand (Nusche et al. 2012), may allow teachers to use their discretion in choosing the tests that will target what they need to assess to help them make the best possible decision regarding what should be the next steps in students' learning.

In a major international study that compares the educational assessment frameworks of some of its member states, OECD (2013) emphasized the main policy directions to help develop 'synergies for better learning.' One of the most striking characteristics of these policy directions is their high degree of interconnectedness. Although they are not solely meant to apply to AfL implementation, they help emphasize the fact that AfL implementation simultaneously involves several targets and that synergies with other policy directions are needed. For instance, 'ensuring a good balance between formative and summative assessment' is more likely to be achieved if there are 'safeguards against an overreliance on standardized assessment' and if a 'variety of assessment types' are used (OECD 2013, Table 2, pp. 21–22). The OECD report (2013) warns:

Not all of the policy directions apply equally to all countries. In a number of cases many, or most, of the policy suggestions are already in place, while for other countries they may have less relevance because of different social, economic and educational structures and traditions. This is a challenging agenda, but tackling one area without appropriate policy attention to inter-related aspects will lead to only partial results. Nevertheless, it is difficult to address all areas simultaneously, and resource constraints mean that trade-offs are inevitable (p. 21).

One major policy direction resulting from the OECD report recommends creating an environment that allows policy targets to be optimally achieved regardless of the constraints. At this juncture point, developing synergies to implement AfL is not then about *controlling or regulating the environment*, it is about *allowing co-regulations to occur*. The main policy directions of the OECD report show a high degree of flexibility and allow for the use of managerial as well as cognitive discretion in several cases. For instance, one of the main policy directions insists on promoting national consistency while making space for local diversity. A level of flexibility can also be found in the reassertion of the crucial role of teacher-based assessment and on the importance of promoting teacher professionalism.

One other important remark that can be made from the report (OECD 2013) is that the main policy directions involve improving assessment sources—'ensure consistency of assessment'—as well as assessors' skills through capacity building—'sustain efforts to improve capacity for assessment and evaluation' and 'build students' capacity to engage in their own assessment' (p. 22). The direction also regards aligning sources of information and human resources with educational goals and students' learning objectives.

Hence, main policy directions require that we focus on three large categories of actions:

1. *Improving the sources of information* by helping, for instance, teachers to create an interface between the information generated by classroom assessment and information from externally designed assessments.
2. *Improving the capacity of assessors* (through skills and capacity building and professional development) at all levels of a school system: for instance, professionals and school leaders in charge of designing and implementing policy decisions, local jurisdiction school leaders, and head teachers and teachers.
3. *Alignment*. In most instances, alignment refers to ensuring that important contents of a program of study are attended to by teachers and that students have been provided with sufficient opportunities to learn them. Alignment can easily be extended to teaching and assessment frameworks, as noted in the Glossary of Education Reform (Coherent curriculum 2014): ‘it [alignment] also refers to coherence among all the many elements that are entailed in educating students, including assessments, standardized tests, textbooks, assignments, lessons, and instructional techniques.’

## 2.6 New Perspectives on AfL Policy Implementation

The OECD study reported above (OECD 2013) stresses the importance of developing synergies among policy directions, and all contributions in Part I of this book are consistent with this view. However, they would all insist, for each contribution, on the need for policies to be adaptable:

- To disruptions introduced by policy changes (Chap. 3—Adie and Willis)
- To the needs of special education students (Chap. 4—Cumming and van der Kleij)
- To where teachers are in terms of assessment literacy and teaching competence (Chap. 5—Griffin et al.)
- To the national and local contexts when policy is a direct import from another international jurisdiction (Chap. 6—Poskitt)
- To time constraints and limitations regarding teachers’ capacity to attend to increased demands in both summative and formative assessment (Chap. 7—Spencer and Hayward)

These items all share the optimistic view expressed by Adie and Willis regarding policy disruptions and overlaps: all adaptation challenges listed previously may be considered as ‘opportunities for professional conversations and changes to pedagogies and assessment practices’ (Chap. 3—Adie and Willis).

The crucial role of teachers’ professional conversations is latent in nearly all contributions. In the specific case of understanding Australia national standards, having teachers work collaboratively made it possible to have AfL policy work hand in hand with curriculum policy enactment (Chap. 3—Adie and Willis). The teachers helped policy makers integrate different policies and explain how policies, which may initially be considered as unrelated by teachers, may fit together.

Working with colleagues helps teachers become more aware of different ways of interpreting the policy documents and be more self-critical of their own interpretation. Working together also assists teachers in developing their professional judgement because policies are not always clear, and decisions must be made regarding the meaning that will be accentuated and used. In the specific cases of understanding the New Zealand national standards, the knowledge acquired by teachers was necessary to enable them to inform students properly on what was expected in terms of achievements. Teachers' mutual understanding of standards helped to improve teacher–student communication on assessment criteria.

In the case of students with disabilities or special needs, teacher–student communication may require a distinctive application of AfL generic principles. Cumming and van der Kleij stress the importance of closely focussing on the manner in which Australian AfL policy is implemented with those students for whom AfL is most likely to be useful and necessary (Chap. 4—Cumming and van der Kleij). For instance, one of the definitions of AfL in Sect. 1.2 of Chap. 1 made reference to teacher–student *dialogue*. AfL practices make intensive use of spoken language. This overreliance on language communication may not be appropriate with students who, for whatever reasons, have major language disabilities.

AfL needs to be adapted to become an instrument of an equity policy for student learners with disabilities (Chap. 4—Cumming and van der Kleij). The equity dimension of any AfL policy enactment should stress the need to enforce the policy differently to take into account not only regional realities but also the characteristics of students with disabilities. Differentiation in how teachers apply AfL generic principles is necessary in order for AfL to be of service to all students.

While differentiation in AfL is needed for special education children, a form of differentiation is required on the part of teachers who need to learn to use AfL with their students. Griffin et al. emphasize that policy implementation should be based on a rigorous analysis of where the teachers are in terms of professional learning and on what the next steps should be for the majority of them (Chap. 5—Griffin et al.). Teachers' professional learning should be based on an AfL learning progression, some knowledge and skills being prerequisites to others. In their analysis of the Philippine situation, Griffin et al. recommended the following first step: 'Teachers need to be supported in framing questions both for assessment and for teaching purposes.... This change alone would have an important impact on the use of formative assessment and would blend assessment with teaching' (Chap. 5—Griffin et al.). This recommendation regards factoring in the teachers' zone of proximal development in the enactment of an AfL policy. This recommendation suggests that any jurisdiction considering designing and implementing AfL on a large scale should begin with a thorough study of the general degree of teachers' preparedness before planning for change.

Differentiation and adaptation also need to occur at the 'policy adaptation' level. The access to multiple foreign experiences in AfL policy enactment has provided researchers and educational leaders with numerous opportunities to learn best policy designs and implementation practices around the world. Poskitt warns that these practices cannot be transferred as they are. Her contribution focuses our

attention on the need to involve all parties and stakeholders in policy adaptation and policy implementation because the basic thinking behind a policy that may have occurred elsewhere still needs to be performed anew if the same policy is to be *adopted* and *adapted* in another jurisdiction (Chap. 6—Poskitt).

When considering other countries as sources of information on AfL implementation, Scotland is often considered as one of the most experienced national jurisdictions on the matter. Spencer and Hayward discuss the lessons learned from the first *wave* of AfL implementation (Assessment *is* for Learning) and the challenges that have emerged as Scotland undertakes a series of major changes both on curriculum (*Curriculum for Excellence*) and on assessment (*Assessment at Transition*) (Chap. 7—Spencer and Hayward).

The challenges met in Scotland went well beyond the issue of implementation. These challenges concerned sustaining the achievements of a first wave of a successful implementation and coordinating existing practices with the requirements of a new curriculum and assessment policy. It simply could not be assumed that a natural integration would occur. To make AfL implementation sustainable required that policy and practice be in close alignment. Research played an important role in providing the evidence base required to realign policy and practice and inform future actions: ‘A major challenge for Curriculum for Excellence was to merge the new ideas about curriculum and learning processes with the preexisting successful assessment for learning practice’ (Chap. 7—Spencer and Hayward).

Such important changes and coordination of policies with existing practices need to be planned for. Improvement in teachers guidance as well as allowing space for co-regulation to occur between different levels of the educational system are research-based decisions that were determined by Scotland education authorities:

Using research to explore the interrelationship of policy and practice as an evidence base to inform future action can help to realign policy aspirations and practice in schools and classrooms. Action based on evidence is the only way to build education systems that are truly learning systems (Chap. 7—Spencer and Hayward).

Synergies need to be developed between curriculum, pedagogy, and assessment as a coherent whole (Wyse et al. 2016). One cannot simply expect such synergies will occur by themselves as if they were self-evident. Teachers and actors in the education system must have opportunities to share their understanding of the policy targets and actions (Chap. 3—Adie and Willis). The Scottish experience also reminds us that policies should set targets at a level that is suitable with teachers’ existing assessment literacy and AfL competence (Chap. 7—Spencer and Hayward). To meet the challenge of implementation, both the Philippines first-time experience in implementing AfL (Chap. 5—Griffin et al.) and the Scottish long-term experience in making AfL sustainable and coherent (Chap. 7—Spencer and Hayward) indicate that policies must be informed by research evidence and a rigorous analysis of where the teachers are in terms of professional learning.

## 2.7 Moving Forward

The enactment of AfL policy in education systems is very demanding. This enactment involves a series of important adaptations not only on the part of teachers but also on the part of the policy designers and the school leaders in charge of implementing those policies. Such adaptations require planning for co-regulations to occur and develop synergies. Time for professional learning is a rare resource. Developing synergies extends well beyond having teachers work together; it also regards educators learning to work together *efficiently* in a manner that is profitable, both collectively and individually. The contributions of Part I demonstrate that synergies are desirable and possible.

Poskitt illustrates how time invested in communication and in involving teachers and other stakeholders is thereafter repaid (Chap. 6—Poskitt). She explains how politics and policies have interacted in New Zealand for the best and for the worst. This raises several issues. Are communication and collaboration and participation in policy processes always successful? Are there situations in which urgent matters would justify that such processes be skipped? In policy enactment, as in politics, consensus is rarely obtained, and there will always be resistance to change. This observation raises yet another question: how is resistance to change addressed in collaborative approaches compared to top-down approaches? Is there space for minority opinions, and if so, are they considered?

Resistance to change may occur for certain appropriate and legitimate reasons. There are limitations to the capacity of teachers and of the entire educational system to assimilate new trends and to accommodate existing practices to changing conditions while maintaining a certain degree of coherence. We simply cannot assume that such capacities already exist, that they can be acquired rapidly, or that such changes would not have an undesirable impact on already existing capacities or practices. Spencer and Hayward warn against the danger of considering teachers as professionals while not providing sufficient time for their professional development (Chap. 7—Spencer and Hayward).

This statement means that, more than ever before, policy implementation must take advance notice of where teachers are in their professional learning (Chap. 5—Griffin et al.). This statement also raises important questions regarding what skills and capacities need to be developed: What should be a teacher's learning progression in AfL? Are there certain necessary steps or prerequisites that would invariably be the same, which resemble developmental stages or standards of progression?

The enactment of AfL policy must also compete with the enactment of other policies. For instance, while teachers assimilate general principles of AfL, they need to accommodate these principles to consider the different needs of special students (Chap. 4—Cumming and van der Kleij). Adding new tasks to existing ones may also put teachers in a state of cognitive dissonance when they encounter what are apparently competing demands of their time and efforts. For instance, in Scotland, the pressure on teachers and students for more frequent summative assessment

diverted time and efforts from AfL (Chap. 7—Spencer and Hayward). Although new demands for assessment *of* learning were not intended to compete with the demands for assessment *for* learning, the time needed on the part of teachers to comply with both actually put teachers in an uneasy situation. The Scottish experience raises several important questions: What are realistic assessment demands? Can we really prioritize AfL over the assessment of learning (AoL), and is there a point where AoL requirements may be preventing AfL from truly being implemented and sustained? According to Spencer and Hayward, ‘it is important to prioritise assessment activities, a process that entails stopping doing some things in order to make it possible to do other, more desirable, things well’ (Chap. 7—Spencer and Hayward).

While a better coordination of efforts between AoL and AfL is necessary for successful policy enactment, AfL cannot play its important role without a similar coordination of efforts with curriculum development. Because AfL also regards helping teachers make the best possible decision regarding what should be the next step in learning for students—which is a most difficult task for teachers as we will observe in Chap. 8—teaching and assessment need to be properly aligned with the curriculum. As shown in Adie and Willis’ contribution, ‘reconciling standards-referenced curriculum and assessment with improved teaching and learning practices necessitates that policy makers also take up the unifying narrative of AfL and reflect this in policy documents’ (Chap. 3—Adie and Willis).

Confronted with the challenges of AfL implementation, there are reasonable grounds for optimism. Even though all Part I contributions illustrated certain shortcomings of the policies and of their implementation, they all provided certain practical solutions we can use to make recommendations.

The metaphor of ‘expansion joint’ or ‘movement joint’ may help illustrate the important role of a *co-regulation space* to allow periodical adjustments to be made in the implementation and enactment of AfL policy. Expansion joints are used to imbue sufficient flexibility in the structure of a bridge so it can adapt to changing climate conditions. Similarly, *co-regulation juncture points* are necessary to absorb the stress generated by the requirements of an education system for change and adaptation. The more challenging the implementation, the more pressure on the system and the more important co-regulation spaces become.

One immediate recommendation that originates from Part I is that policy should plan for co-regulation spaces in the AfL implementation process. Such co-regulation spaces allow for preventive and early adjustments should something unexpected occur or go wrong. The spaces also provide policy designers and enactors with margins of tolerance and discretion to adjust and adapt. The sooner the trajectory of policy enactment can be corrected, the least effort will be needed to readjust the target or the trajectory and the least frustration that will occur among enactors. Poskitt provides a very suitable illustration of two diametrically opposed policy implementations in New Zealand (Chap. 6—Poskitt).

Successful AfL policy implementation requires that co-regulation spaces be planned early not only to prevent or correct the misalignment of policy enactment but also to afford opportunities to develop the capacity and synergies needed to

improve efficiency and save time and energy on a continuous basis. It is clear from Part I contributions that the time invested in communication and involving teachers and other stakeholders is thereafter repaid. Here is a short list of recommendations that can be deduced from Part I contributions to move forward on meeting the challenges of implementation:

1. Begin with a thorough study of the general degree of teachers' preparedness before planning for change. Consider teachers' zones of proximal development in the enactment of an AfL policy (Chap. 5—Griffin et al.).
2. Provide opportunities for teachers' collaborative work to help teachers become more aware of the different means of interpreting the policy documents and be more self-critical of their own interpretation (Chap. 3—Adie and Willis).
3. Involve all parties and stakeholders in policy adaptation and policy implementation (Chap. 6—Poskitt). Teachers' mutual understanding of standards helps to improve teacher–student communication regarding the assessment criteria.
4. Enforce the policy differently to consider not only regional realities but also the characteristics of students with disabilities (Chap. 4—Cumming and van der Kleij).
5. Use research evidence to realign policy and practice and inform future actions. Plan for the coordination of policies with existing practices (Chap. 7—Spencer and Hayward).

Whatever the policy and its merits, the previous recommendations are not realizable if certain efforts are not directed at developing the required capacity to properly enact policy objectives through the professional development and collaborative learning of teachers and other stakeholders. Professional development and collaborative learning are essential components of policy enactment and will be the topic of Part II, *Building Capacity: Professional Development and Collaborative Learning about Assessment*.

## References

- Ben Jaafar, S., & Anderson, S. (2007). Policy trends and tensions in accountability for educational management and services in Canada. *The Alberta Journal of Educational Research*, 53(2), 207–227.
- Berman, P., & McLaughlin, W. M. (1978). *Federal programs supporting educational change: Vol. VIII*. Santa Monica, California: Implementing and sustaining innovations.
- Black, P., & Wiliam, D. (2003). In praise of educational research: Formative assessment. *British Educational Research Journal*, 29, 623–637.
- Blackmore, J. (1988). *Assessment and accountability*. Victoria, Australia: Deakin University.
- Christie, C. A., & Fierro, L. A. (2012). Evaluation policy to implementation: An evaluation of scientifically based research in practice. *Studies in Educational Evaluation*, 38, 65–72.
- Coherent Curriculum. (2014, March 3). In S. Abbott (Ed.), *The glossary of education reform*. <http://edglossary.org/coherent-curriculum/>. Accessed 5 Nov 2015.



- Earl, L., & Timperley, H. (2014). Challenging conceptions of assessment. In C. Wyatt-Smith, V. Klenowski & P. Colbert (Eds.), *Designing assessment for quality learning*. Dordrecht: Springer.
- Halverson, R. R., & Clifford, M. (2006). Evaluation in the wild: A distributed cognition perspective on teacher assessment. *Educational Administration Quarterly*, 42(4), 578–619.
- Hodson, R., & Sullivan, T. A. (2012). *The social organization of work* (5th ed.). Belmont, California: Wadsworth.
- Nusche, D., Laveault, D., MacBeath, J., & Santiago, P. (2012). *OECD Reviews of evaluation and assessment in education: New Zealand 2011*. Paris: OECD Publishing. doi:10.1787/9789264116917~en. Accessed 5 Nov 2015.
- OECD. (2011). *Evaluation and assessment frameworks for improving school outcomes. Common policy challenges*. Paris: OECD Publishing. <http://www.oecd.org/education/school/46927511.pdf>. Accessed 20 Sept 2014.
- OECD. (2013). *Synergies for better learning: An international perspective on evaluation and assessment*. OECD reviews of evaluation and assessment in education. Paris: OECD.
- Seifert, C. M., & Hutchins, E. L. (1992). Error as opportunity: Learning in a cooperative task. *Human Computer Interaction*, 7, 409–435.
- Trochim, W. M. K. (2009). Evaluation policy and evaluation practice. In W. M. K. Trochim, M. M. Mark & L. J. Cooksy (Eds.), *Evaluation policy and evaluation practice. New Directions for Evaluation* (Vol. 123, pp. 13–32).
- Weinbaum, E. H., & Supovitz, J. A. (2010). Planning ahead: Make program implementation more predictable. *Phi Delta Kappan*, 91(7), 68–71.
- Wyse, D., Hayward, L., & Pandya, J. (Eds.). (2016). *The SAGE handbook of curriculum, pedagogy and assessment*. London: Sage Publications.

# Chapter 3

## Making Meaning of Assessment Policy in Australia Through Teacher Assessment Conversations

Lenore Adie and Jill Willis

**Abstract** In Queensland, Australia, despite a long tradition of classroom-based assessment and scholarly assessment research, assessment for learning (AfL) has not had a distinct identity as a set of ideas within daily teacher classroom practice. Rather, the initial research by the Assessment Reform Group that sparked reform in other contexts has been accommodated into existing policies and practices. This has resulted in missed opportunities for teachers to engage in deep inquiry into the underpinning and interconnected philosophy of AfL as a suite of practices that inform ongoing teacher and student dialogue into improving learning. However, recent national assessment policy changes have disrupted curriculum planning, assessment and reporting practices and enabled renewed conversations about the role of assessment in informing classroom learning. This chapter focuses on the first phase of AfL classroom practice that involved developing shared teacher understanding of assessment standards. We suggest that this dialogue about standards at the beginning of the teaching semester is a necessary precursor to informed teaching that involves the sharing of expected standards with students, and is an opportunity for teachers to engage with the philosophy of AfL.

### 3.1 Introduction

Assessment for learning (AfL) policy has a tenuous and tentative relationship with enacted assessment practices in Australia and specifically, in Queensland. The system of assessment in Queensland has been based on the professional assessment capacity of teachers for over forty years, and may be considered as an ideal context within which to develop effective AfL practices. However, without the historical tension of external examinations that seems to have provoked an impetus for AfL in

---

L. Adie (✉)  
Australian Catholic University, Brisbane, Australia  
e-mail: Lenore.Adie@acu.edu.au

J. Willis  
Queensland University of Technology, Brisbane, Australia

other nations, AfL has remained a minor policy theme in Queensland. AfL as a policy imperative in Australian schools, and Queensland schools in particular, is not as clearly defined as it has been in other policy contexts such as in England, Scotland, and New Zealand. This means that as an identified practice, AfL has not had a coherent narrative within the education and assessment policy that informs teachers' work in Queensland schools.

AfL is understood in this chapter as a philosophy of interconnected pedagogic, curriculum and assessment practices that challenges teachers to include students as co-constructors of meaning about their learning journeys (Earl and Timperley 2014; Popham 2014). As Stiggins (2002) noted, 'Assessment *for learning* is about far more than testing more frequently or providing teachers with evidence so that they can revise instruction, although these steps are part of it...we now understand assessment *for learning* must involve students in the process' (p. 761). Yet, internationally it has been reported that the transformative power of AfL has been misrepresented or misunderstood in policy agendas with students positioned as subjects of assessment policy instead of owners of their learning (Looney 2014; Swaffield 2011). Concerned that AfL policy and practices were emphasising a superficial understanding of strategies, participants at the Third International Conference on Assessment for Learning proposed a second generation definition of AfL in 2009. These international experts noted that: 'Assessment for Learning is part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information from dialogue, demonstration and observation in ways that enhance ongoing learning' (Klenowski 2009, p. 264).

When teachers implement AfL by introducing strategies (that is, the letter of AfL) without a conceptual or philosophical understanding of AfL (the spirit of AfL), the potential for learners to develop greater autonomy in their learning is severely constrained (Marshall and Drummond 2006; Willis 2011). Earl and Timperley (2014) argue that to get beyond tinkering and a policy focus on the prepositions 'of,' 'for,' and 'as' assessment, teachers need to engage in deep professional learning that can lead to conceptual change. Teachers who adopt the spirit of AfL are driven to provide opportunities for students to learn, with every curriculum activity designed to give teachers insight into what students are thinking and for students to have the knowledge to decide what to do next. Pryor and Crossouard (2008) describe this process of meaning making about assessment quality as a social practice that occurs through action and over time as a collaborative narration of identity. This focus on continual learning is often a paradigm shift for teachers that requires intentional rethinking, sustained attention, and energy (Earl and Timperley 2014).

The introduction in Australia of a national curriculum and associated year level achievement standards, as well as national professional standards has started to disrupt previously held narratives of assessment and learning, and has been a potential impetus for teachers to re-examine their assessment practices. This chapter presents our reflections about implementing the spirit of AfL into Queensland schools after four years of working alongside teachers in these changing policy contexts. The chapter commences with a consideration of policy as narrative,

followed by a brief history of assessment policy in Queensland, and an overview of the Australian Professional Standards for Teachers in order to contextualise the discussion. Conversations with teachers as they negotiated the meaning of a new national curriculum and achievement standards into their existing classroom assessment practices are drawn from to illustrate the challenge of implementing AfL in a particular policy context, and how this influences the enacted practice.

### 3.2 Policy as Narratives of Meaning

Policies can be narratives of meaning that capture a grand social vision for society, as ‘epic poems or stories...about durable historical, social and cultural commitments to particular forms of education’ (Luke 2011, p. 374). When assessment principles underpin the assessment policy, the policy narratives can be viewed as collaborative responses between policymakers, researchers, and practitioners, with policy providing support and resources to teachers and schools to enact assessment (Griffin 2007). More frequently, policy positions schools as problems in need of reform, and policy becomes a type of script that authorises what counts as a desirable performance (Fitzgerald and Savage 2013). In a process of global policy borrowing where policies cross national borders at ‘unprecedented scale and speed’ there is a risk that instead of educational policies being long-standing settlements arising from social and cultural debate, policies can become flattened and limited in meaning (Luke 2011, p. 367). There has been a loss of meaning as AfL has been adopted as part of a global assessment policy movement that Looney (2014) argues is ‘both viral and normative’ (p. 234). Looney draws on evidence across several nations to show that AfL has often been transformed from a narrative of good news for learning into a policy narrative associated with performativity and managerialism. Likewise, Black (2007) has described the prolific use of the term assessment for learning as ‘superficial adoption’ (p. 18) which has detracted attention from the educational goal of improving student learning.

Policy is not simply read or received by schools but is understood as a social and constructive process of meaning making by teachers and students. Bernstein (2003) provides a theoretical language to describe the translation of policy by teachers into pedagogic practice, and how this translation intersects with culturally specific local practices as recontextualisation. When significant changes to national curriculum and assessment occur, specialist knowledge becomes translated into accessible forms or ‘pedagogised to constitute school knowledge’ within local contexts (Singh 2002, p. 571). This involves knowledge agents such as federal and state education systems, policy writers, local school authorities, textbook writers, teacher educators, teachers, parents, and students recontextualising the knowledge, as each person ‘selectively appropriates, relocates, refocuses, and relates other discourses to constitute its own order and orderings’ (Bernstein 2003, p. 159). Recontextualisation changes how the meaning of the knowledge is represented by the agent, such as a teacher, to others such as colleagues and students, as the agent decides what is more

or less important, and how to relate the new knowledge, in this case assessment standards, within daily practice. In a similar manner, students through their dialogue with their teachers and their peers shape meaning for themselves. Meaning making is a social process as teachers and students negotiate, interpret, and reinterpret the expected assessment practices as directed through the Australian Curriculum and other policy narratives. It is with this understanding that the authors focus on the importance of assessment conversations as opportunities to make meaning from policy and other related texts.

One of the significant challenges of AfL implementation is understanding how teachers engage in the complex conceptual work of shifting assessment paradigms towards AfL principles, while recontextualising and navigating the meanings of disconnected assessment policies. In Australia, the language of AfL has been assimilated into various national and state assessment policy discourses without the accompanying cultural debate. AfL was initially promoted through a website in Australia hosted by the Curriculum Corporation (now Education Services Australia<sup>1</sup>), a national curriculum service provider governed by representatives of the state ministers for education. The AfL definition from the Assessment Reform Group (ARG) appears on the opening page stating that AfL is ‘the process of seeking and interpreting evidence for use by learners and teachers to decide where the learners are in their learning, where they need to go and how best to get there’ (Assessment Reform Group 2002, p. 2). Included on the website are a range of support materials, sample assessment tasks and responses, and professional development materials. With the development of the site around 2010, DVDs of classroom practice were distributed to all schools. In Queensland, this federal resource was not incorporated into any state assessment policies, and so AfL principles became part of the ‘collection code of unconnected bits and pieces’ of policy in schools, with school leaders and teachers taking on the policy actor roles of narrators and translators (Ball et al. 2011, p. 627). Teachers’ work and the positioning of AfL practices within Queensland have been influenced by the state’s history of assessment and other recently introduced policy frameworks, including a national standards-referenced curriculum and national professional standards for teachers.

### 3.3 The History of Assessment Policy in Queensland

Queensland has had a system of school-based assessment and teacher professional judgement since 1972 (Queensland Curriculum and Assessment Authority 2014). The freedom for teachers to develop their own assessment tasks might be highly congruent with AfL, yet research in a large sample of Queensland classrooms

---

<sup>1</sup>This information is accessible through <http://www.assessmentforlearning.edu.au/default.asp?id=912>.

confirmed that assessment tasks were generally busywork, reproductive, shallow, and underestimating of student ability (Lingard et al. 2006). Since 2008, there has been an increasing emphasis in state and federal policy on assessment task standardisation through a variety of assessment tools including the introduction of Queensland Common Assessment Tasks (QCATs), and the introduction of National literacy and numeracy tests (NAPLAN) for Years 3, 5, 7, and 9 (for children aged 8, 10, 12, and 14 years). These assessment innovations have significantly influenced teacher practice, yet the dominant policy narrative was often one of consistency of judgement and teacher improvement rather than a focus on students as agents in their learning journey.

QCATs for Years 4, 6, and 9 (for children aged 9, 11, and 14 years) were introduced into Queensland schools in 2009 after an extensive consultation process. The purpose was to build teacher professional knowledge through modelling a task design that could evaluate higher order thinking skills (Freebody 2005), and to develop shared understanding of the relevant year and subject standards through a social moderation process (Queensland Curriculum and Assessment Authority Curriculum and 2015). Over the course of their implementation, teachers developed a shared language of assessment and knowledge of assessment design, yet the focus remained on teacher learning rather than a conceptual shift for using assessment to inform student learning. While the tasks are still available for teachers to use and draw from in their own assessments, production of new tasks ceased in 2012 (Queensland Curriculum and Assessment Authority 2015). This change coincided with a change of government in Queensland, the introduction of a national curriculum, and the increasing influence of national assessment.

A National Assessment Program—Literacy and Numeracy (NAPLAN)—was introduced into Australian schools in 2008. These tests in reading, writing, language conventions, and numeracy occur in the second week in May in every school throughout Australia for approximately 1 million students (Australian Curriculum, Assessment and Reporting Authority 2013a). The stated policy narrative is that schools and teachers will use this data to develop and strengthen their teaching and learning programs such that every student progresses in their learning, that is, high achievers are continually challenged, and students requiring support are identified (Australian Curriculum, Assessment and Reporting Authority 2013b). While this official policy statement might be congruent with AfL intentions, the policy impact has been quite different.

Initially the media reports of NAPLAN results were presented as a comparison between overall state and territory results, which evoked a sense of competition rather than an emphasis on formative learning. Queensland was placed low amongst the eight states and territories across all domains of reading, writing, and numeracy. The political and public perception of the results of Queensland children in the first national exam performance had been so distressing that it led to an investigation and policy recommendations that included preservice and in-service teacher professional learning (Masters 2009). In January 2010, the test results for individual schools were released publically through the media. This occurred amid warnings from international experts that this was a fraught process that could result in

misrepresentation of the data (Berliner 2010; Goldstein 2010). Teacher assessment and classroom practice in Queensland has since been the subject of intense public and political focus.

Queensland classrooms have moved from being privatised to deprivatised sites of practice quite rapidly, with the driving narrative being one of accountability, performativity, and measurement. The NAPLAN results provided an impetus for several actions:

- a greater focus on pedagogical frameworks within schools;
- a move for Year 7 students (aged 12 years) from the primary sector (children aged 5–11 years) to the secondary sector (children aged 12–17 years);
- an increased accountability of principals (and teachers) for results; and
- an increasing expectation of informed practice through professional learning.

Queensland's perceived poor performance also triggered a review of the Senior Assessment system that was based on continuous teacher-created assessments, with folios moderated by teacher panels; the quality of graduates being attracted to the teaching profession; the quality of teacher education; and a restructuring of the state's curriculum and assessment authority. The review into Queensland's senior assessment and tertiary entrance systems (Matters and Masters 2014) resulted in key recommendations and a proposed redesign of the assessment system to include external assessments that are set and marked by the curriculum and assessment authority, a move away from Queensland's quite unique focus on school-based assessment.

Interwoven with Queensland's historical narrative of school-based assessment and the introduction of national standardised testing regimes, has been the shift to a national curriculum since 2010, as well as the introduction of national professional standards in 2013. The response by state education systems to both of these events is significant to our story of AfL in Queensland. Connected to the national curriculum is an expanding set of related resources, as well as the state specific guidance from the Queensland Curriculum and Assessment Authority (QCAA). However, references to AfL within the Australian curriculum documents are difficult to find. A search through the online national curriculum documents for the term 'assessment for learning' returned no results. A generic statement identifies that one of the recommended approaches to assessment is the inclusion of 'ongoing formative assessment within classrooms for the purposes of monitoring learning and providing feedback, to teachers to inform their teaching, and for students to inform their learning' (Australian Curriculum, Assessment and Reporting Authority 2015). Teachers who search through the linked resources will find tools related to feedback and peer and self-assessment but not a holistic narrative of assessment for learning as can be found in the policy discourse of other nations. For example, in New Zealand, assessment is described as a process of learning, for learning (New Zealand Ministry of Education 2011, p. 12), and positions students to be 'educated in ways that develop their capacity to assess their own learning' (Absolum et al. 2009). In Scotland, assessment policy encourages 'dialogue about learning,'

between teachers and students who together ‘identify next steps and learning goals based on feedback and evidence of learning’ (Education Scotland 2015). While we are not suggesting that these elements are missing in the teaching practices of Queensland teachers, the explicit discourse is hidden in Australian policy amongst an imperative for teachers to evidence learning.

Australian and Queensland teachers are also guided by a set of seven Professional Standards which outline ‘what teachers should know and be able to do’ (Australian Institute for Teaching School Leadership 2014, paragraph 1). The standards were introduced into Queensland schools in 2012. Professional standards were not new for Queensland teachers who, since 2006, had been using these for registration purposes. However, the national professional standards came at a time of increasing accountability, performativity, and evidencing of quality for teachers and principals (Tuinamuana 2011). Across the professional standards, teachers are expected to use assessment to inform curriculum planning (2.3), establish learning goals (3.1), use feedback and student assessment results to inform planning and improve programs (3.6), contribute to collegial discussions (6.3), and continue their own professional learning to improve student learning outcomes (6.4). Standard 5 is devoted to teacher assessment practices and notes that amongst other assessment practices a highly accomplished teacher will diagnose learning needs, provide targeted feedback, use judgements to progress student learning, develop consistent judgements through moderation, use data to identify interventions, and construct accurate reports to students and parents (Australian Institute for Teaching School Leadership 2014). While many underpinning elements of AfL can be individually recognised within the professional standards, the focus within the standards is once more about the work of teachers as diagnosticians providing assessment information ‘to’ students and parents rather than involving students in a dialogue about learning. The potential for the Australian Professional Standards for Teachers to inform teacher adaptive expertise in AfL is not high, as the AfL narrative remains disjointed, patchy, and teacher-centred.

More recently, teachers in Queensland have been reintroduced to the language of AfL strategies, such as learning intentions and success criteria, and effective feedback, through pedagogic models such as Visible Learning (Hattie 2009), the Art and Science of Teaching (Marzano 2007), Direct Teaching (Fleming and Kleinhenz 2007), and AfL (especially in the Catholic school system). The work of John Hattie (2009) has been particularly influential in Queensland education policy, with a heightened focus on assessment that improves learning. As this rethinking occurs, it has opened up opportunities to once again discuss both formative and summative assessment practices within curriculum design and pedagogy. Quality teaching practices, including those advocated as AfL are evident in many schools, however there is not an explicit policy narrative of AfL. Policy advice to reposition assessment as a co-construction between teacher and student and a ‘continual flow of information about student achievement... in order to advance, not just check on, student learning’ drawn from the work of Rick Stiggins, is hidden deep inside websites in teacher support materials (Queensland Studies Authority n.d.). In this chapter we suggest that the policy patchwork has put the burden of



recontextualising multiple policies into pedagogic practice onto teachers, without the helpful guidance of an underpinning and coherent conceptualisation of the purposes of AfL.

### 3.4 Research Context

In the rest of this chapter, we interweave examples from a four-year research project to illustrate and analyse both the practical and theoretical complexities that were occurring for teachers as assessment policy was recontextualised into teaching practice and student learning. We consider the different ways that policy has disrupted teacher practice and how the teachers' conversations helped them to connect the curriculum, assessment, and pedagogic narratives. This was a process of recontextualising policy into practice. Each of these stages or layers required the teachers to be policy actors (Ball et al. 2011) who were actively negotiating meaning. In the final section of this chapter we present a representation of the importance of teachers developing a shared understanding of standards as an early phase of AfL. We suggest that when this development of shared understanding occurs at the beginning of a teaching semester, it enables a flow into teaching practice that informs a purposeful and planned use of assessment for learning strategies.

This research project took place from 2011, soon after the Australian Curriculum documents had been produced and released online, to 2014 as teachers started to implement the curriculum into their classrooms. Since this time the authors have worked with and presented to a range of school sectors and teachers on the use of dialogue and annotations to develop shared understandings of the achievement standards. We draw on the data from five primary schools (children aged 5–11 years) and seventeen teachers as they worked within new curriculum contexts and attempted to understand these practices within their historic ways of working and their specific school culture, as well as our reflections from working with state sector authorities and workshops with both primary and secondary teachers. We worked alongside teachers in the five schools in classes ranging from Year 2 to Year 6 (Table 3.1).

In 2011, we worked with nine Year 6 teachers from three schools as they negotiated the meaning of the new Australian Curriculum and the associated end-of-year achievement standard within their practice, and considered how they would use the standard statement to make judgements of student work on a marking scale of A to E (Willis and Adie 2013). Year 6 was chosen as teachers of this year level may have engaged with annotated work samples through QCATs, and it was not a year level in which national testing occurred, so the assessment policy pressures were not as great. In each school, the teachers were released from classes for a whole day of professional conversation, facilitated by the researchers. Detailed notes and recordings of all of the conversations were made. The focus of these conversations was twofold: (1) understanding the implications for planning when

**Table 3.1** Modes of working with teachers

Year	Mode of working	School level	
		Primary	Secondary
2011	Intensive small teaching teams	Year 6 (children aged 11 years) 3 schools; 9 teachers	
2012	Workshops and presentations	Independent and state school leaders and teachers (n ~220)	
2013	Intensive small teaching teams	Year 6, 1 school, 5 teachers and head of curriculum Year 2 (children aged 7 years) 1 school, 2 teachers	
2014	Workshops and presentations	Independent and state school leaders and teachers (n ~350)	

using an end-of-year achievement standard, and (2) interpreting the achievement standard as an A to E judgement, by articulating the qualities that were being valued when making judgements. The teachers were encouraged to develop sample student portfolios for their school, illustrative of A–E standards in English and Mathematics, and annotate the student work samples using the achievement standard in the Australian Curriculum. The Year 6 teachers from each school met together two months later to moderate the portfolios and share their ideas and approaches to gathering and finding evidence.

During 2012 (and again in 2014) we drew on our findings from our intensive work with small groups of teachers, to deliver workshops and presentations to large groups of teachers during professional development days and through conference presentations. In these presentations, we focussed attention on the Australian Curriculum Achievement Standards, and the implication of this for planning, teaching and assessment practices. Teacher discussions, questions and concerns that arose from these presentations contributed to our further thinking and research directions in the following year. In 2013, we worked with volunteer teachers, two Year 2 teachers from one school, and five Year 6 teachers and one Head of Curriculum from another school. Our specific focus that year was to understand whether teachers who collaboratively annotated student work samples at the beginning of the semester, during the planning process, developed a shared understanding of the achievement standard. We encouraged teachers to reflect back to us how this shared understanding influenced their subsequent teaching practice. Each school was visited several times, and included discussions with the school principal regarding school priorities; an initial planning day with the teachers at the beginning of the semester; and follow-up visits to classrooms and email conversations throughout the year (Adie and Willis 2014; Willis and Adie 2014). These discussions were audio recorded and analysed independently by each researcher and then compared. In this chapter we build on these previous analyses to understand how teachers are creating a coherent assessment narrative that includes the spirit of AFL.

### 3.5 Teachers as Policy Agents, Actors and Meaning Makers

For the teachers featured in this project, recontextualising the newly published Achievement Standards in the Australian Curriculum was a lengthy process that involved several stages. Teachers first had to learn how to access and interpret the achievement standards. The Achievement Standards are published as a digital curriculum that is openly accessed and regularly updated, so that teachers, parents, and students can all directly engage with the statements that describe what students are expected to know and do by the end of each year of schooling (Australian Curriculum, Assessment and Reporting Authority 2014). There was an initial process of recontextualising as these standards provided a new assessment discourse for teachers, and they needed to reinterpret historic understandings through comprehending this new text (Willis and Adie 2013). The purpose and the format were not clear in the texts provided to teachers:

What do I need to know? Spell it out for me in very clear structure ... not the lovely flowery descriptors. I have to do too much thinking. Too much of this is open to interpretation. 'For different purposes?' What are the purposes? (Year 6 teacher, School 1)

The teachers commented that it took several readings to make sense of the dense text. Recontextualizing also occurred as the teachers translated the implications of these standards into everyday local assessment discourses of their school as teaching teams met together to develop annotated exemplars (Adie and Willis 2014; Willis and Adie 2013).

Things that have weighed heavily in [our] rubrics aren't weighted heavily in the achievement standards .... It is something I still value as a teacher, but perhaps it shouldn't be in the rubric to such a degree if it is not in the Australian Curriculum. (Year 6 teacher, School 1)

Working with a new curriculum required the teachers to forego previously cherished content and pedagogic priorities that did not reconcile with the new text of a year level achievement standard. These conversations caused the teachers to deeply analyse this new text and prompted several discussions about valued qualities and ways of working that included the connections between assessment and student learning.

In the first year of this project, working alongside the teachers on the professional learning day, we had analysed annotated exemplars available on the Australian Curriculum website and worked together to translate the new assessment standard into the current practice of using A to E standard descriptors. We had anticipated that, after the day of professional learning, the teachers would then start to collect samples of student work, adding their own annotations that would contribute to portfolios of evidence of a standard. When we met together several weeks later to moderate the portfolios, we found that the teachers had difficulty translating these understandings to different samples of student work and were hesitant to record these as annotations.

In 2013, we worked purposefully with the teachers to annotate work samples in the beginning planning meeting so that both a shared understanding of A to E qualities developed amongst the teaching team and the process of annotating work samples was begun. Annotations were understood as a representation of how the teacher translated the knowledge embedded in the national assessment standards, and in the local school assessment policies, and traditions, into their daily practices. An example of an annotated work sample is included in Fig. 3.1. Collaboratively annotating samples of assessment was challenging work for teachers. It involved prioritising time for intensive discussions with colleagues and depended on trust between colleagues as individual assessment practices were deprivatised.

I was thinking before when we were writing down our thoughts [annotations], what a waste of time, as if we don't all do this [make judgements in similar ways]. And then I realised actually we are all doing very different things. (Year 6 teacher, School 2)

It involved preparedness to work through periods of incoherence and disorder as the teachers had to search for language to articulate knowledge that had previously been tacit and negotiate the implications within school assessment cultures (Willis and Adie 2013). It also involved epistemic tensions such as re-examining beliefs about the positioning of assessment in the teaching cycle and the teacher's identity as assessor (Adie and Willis 2014).

I've developed as a teacher. I'm more aware of the importance of it [mapping backwards from assessment]. (Year 2 teacher, School 3)

The recontextualising process was complex, spanning national, state, and local policies and practices, including; national curriculum, national tests, reporting policies, teacher standards, state assessment practices, concurrent state curriculum documents, and systemic advice about national curriculum implementation and social moderation. Teachers were involved in mapping the new curriculum, and trying to understand new terminology of national achievement standards while simultaneously working within existing school policies about teacher assessment, pedagogic frameworks, reporting processes, and parent expectations. Without a guiding policy narrative this complexity influenced how assessment was conceptualised.

How this deluge of policy, curriculum, and local, state, and national expectations relate to each other is important, and ultimately affects teachers' uptake of practices, such as AfL. If elements configure in a certain way, the opportunity may present for teachers to adopt AfL to inform learning, rather than merely report on learning. For example, annotating student work may be considered as contributing to teacher and student learning while also as part of a quality assurance process. Harlen (2007) describes 'providing samples of work and showing how certain aspects relate to the criteria of assessment' as one form of quality assurance and explains that the focus of this practice 'is on the process of arriving at that outcome' (p. 78). Our work with teachers suggests the importance of teachers developing their understanding of assessment standards within contextualised conversations. Annotated examples of student work have been supplied as support materials for teachers in their

1. Example of teacher’s annotation created during the planning discussion (Extract from Year 4 student’s writing of a fable)

1. Coherence has been achieved through the effective use of pronouns, temporal words (soon, after, after hours, suddenly) to create interest and dramatic tension, and through repetition highlighting Rascal's tricks.

Rascalls favourite trick was hiding, up in the tree branches and jumping down on the other animals heads. This was the worst trick of all. Soon Leo organised a meeting, he was going to do something about Rascals horrible tricks. After hours of thinking Millichad had an Idea “Oooh, I know!” she said. “Let’s play a trick back to Rascal!” “Yes, but what type and how?” questioned Duffy. Suddenly Leo had an Idea. “Aha!” he exclaimed “I know exactly what to do!” So he explained to the other animals.

2. Use of images to create story line –Is “Idea” written with a capital letter to emphasise a “Big Idea”? Author’s voice is apparent in second and final sentences to convey important points in the development of the story line. Range of synonyms used for ‘said’ in text to convey developing drama. Demonstrates understanding and increased control of language to develop plot.

2. Extract from Australian Curriculum Year 4 Achievement Standard with relevant text highlighted (<http://www.australiancurriculum.edu.au/english/curriculum/f-10?layout=1#level4>)

Productive modes (speaking, writing and creating):

Students use language features to create coherence and add detail to their texts. They understand how to express an opinion based on information in a text. They create texts that show understanding of how images and detail can be used to extend key ideas.

Students create structured texts to explain ideas for different audiences. They make presentations and contribute actively to class and group discussions, varying language according to context. They demonstrate understanding of grammar, select vocabulary from a range of resources and use accurate spelling and punctuation, editing their work to improve meaning.

3. Selection of relevant Australian Curriculum Year 4 Content Descriptions  
 Language: Understand how texts are made cohesive through the use of linking devices including pronoun reference and text connectives (ACELA1491)  
 Literature: Create literary texts by developing storylines, characters and settings (ACELT1794)

Fig. 3.1 Example of annotations according to the Australian curriculum achievement standards

moderation practices of the QCATs since 2009 (Queensland Curriculum and Assessment Authority 2015). Our initial assumption in 2011 when we first worked with the teachers was that they would be able to transfer their experience in working with the annotated exemplars and in giving feedback to students to create their own

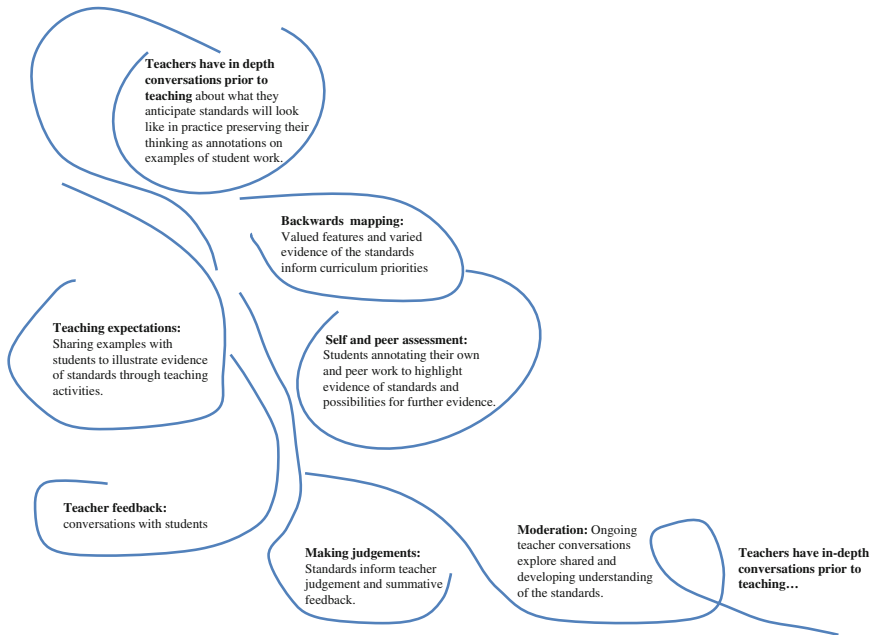
annotations. However, engaging in collaborative annotation prior to curriculum planning emerged as a challenging new purpose for teacher assessment conversations. These conversations enabled teachers to participate in a ‘critical inquiry about the differentiation between achievement standards’ (Adie and Willis 2014, p. 133). Investing time in developing a shared understanding of the achievement standards at the beginning of a semester resulted in a shared and informed teaching practice as teachers used this knowledge to engage their students in coming to know the expected qualities of a standard.

Before teachers can guide students to demonstrate their learning at a given standard, teachers need to be able to clearly articulate the qualities that they are valuing. This understanding of quality, once clear, is then embedded in their teaching practice where students are explicitly taught the elements that make up a skill. This clear explanation becomes part of how the students represent and talk about their learning, which enhances opportunities for self-assessment. This cascading process is illustrated in Fig. 3.2. Our point is that the participation of students in assessment conversations cannot even start if teachers do not first have a clear understanding themselves of the expected standard of learning and the qualities that will evidence this understanding.

Recontextualising policy into teaching practice is a complex process involving interrelated elements that teachers draw together to make meaning for themselves. While we have represented a coherent stream of assessment practice based around shared standards, teachers were simultaneously navigating the rapids of national and state policies, local assessment practices, and personal beliefs. Teachers were also drawing on other pools of knowledge like their understandings about students and their cultural knowledge. The metaphor of a cascade was chosen to represent the connected, turbulent, and fluid process of assessment meaning making that teachers and students navigate. In our observations of teachers negotiating the new practice of annotation, shared reflection through conversation was a necessary process so that teachers could think together about how the elements could begin to flow together.

### **3.6 Meaning Making Through Annotation Conversations**

We observed that teachers needed to invest a significant amount of intellectual and emotional energy into the process of collaboratively annotating student examples, to recontextualise their assessment knowledge and multiple policy agendas into their classroom practice. We agree with Earl and Timperley (2014) that it takes energy and focus to then enact these changes into the classroom with the goal of enabling students to develop a similar understanding for themselves. Our findings indicate that when teachers apply the principles of AfL to their teaching, even young students can rise to meet these expectations.



**Fig. 3.2** Cascading development—standards as a source of AfL shared understanding

I was blown away by the results...who would have believed this was the work of 7 and 8 year olds? ... I didn't think they would be able to achieve this high standard. (Year 2 teacher, School 3)

While some teachers were pleased to have used the annotations only to inform their own learning, some teachers chose to share their annotations with their students to assist them to understand the expected standard and to use this information to self-assess. This flow from teacher dialogue to classroom practice is an example of how teachers used their planning conversations to share their understanding of the expected qualities of performance with their students.

We have been showing the students [annotated] samples from last year and have noticed that this has improved the quality of the work they complete as they can see that either they could do better, or that they can see where to aim for. (Year 6 teacher, School 2)

By supporting students to understand the qualities that provide evidence of a standard, the teachers were empowering their students to take control of their own learning journey. The Australian curriculum and the professional standards for teachers both include an expectation of inclusion of all students in the learning process. Teachers enacting policy imperatives of inclusion naturally start to plan pedagogic strategies that align with the AfL suite of effective strategies so that an understanding of quality can be represented in multiple modes for their students.

Assessment knowledge, or the development of assessment literacies, involves teachers and students developing multiple ways of knowing what quality performances look like and how to make ongoing judgments about those quality performances (Sadler 1998). Assessment literacy for students occurs when the teacher and the students are engaged in a dialogue that involves ‘articulating and negotiating classroom and cultural knowledges...in the initiation, development and practice of assessment to achieve the learning goals of students’ (Willis et al. 2013, p. 242). AfL practices involve both short negotiations of meaning that occur in classroom interactions between teachers and students and the longer term sustained conceptual understanding of a standard that develops through critical inquiry into the qualities that define the differentiated levels of a standard. Negotiating and implementing this deep understanding of the standards is challenging practice for teachers and students, but the goal of improved teaching and learning is worthy of the sustained energy and focus required. Reconciling standards-referenced curriculum and assessment with improved teaching and learning practices necessitates that policymakers also take up a unifying narrative such as AfL and reflect this in policy documents.

### **3.7 Meeting the Challenges for Implementation**

The conclusions of this study are based on our intensive work with 17 teachers and our discussions with many other teachers. These assessment conversations have informed our practical and theoretical thinking about assessment and the following considerations for assessment policy and standards-based reform. We have observed that within education literature, policy, and guidelines, terms such as ‘assessment for learning’ are often used as if their inclusion will provide sanction and authority for the texts, as well as evoke a shared knowledge and understanding of their meaning in practice. What is important for the enactment of the spirit of AfL is embedding an underpinning philosophy of AfL into policy rather than simply using the name, or listing strategies. This would involve teachers understanding the cascading development of coming to know how different performance qualities can be recognised in achievement standards and how this is translated to teaching practice and student learning.

These lessons from Queensland have implications for all systems that introduce new assessment policies. Where there are many policy disruptions and overlaps there are also opportunities for professional conversations and changes to pedagogies and assessment practices. The potential to strengthen teaching practice will only occur if teachers are introduced to the philosophical underpinnings of how assessment can be used to inform learning in relation to the new policy context. Without this foundational understanding AfL becomes the latest fad, ready to be discarded as soon as the next wave of changes arrives. Our belief is that this conceptual understanding of assessment as a shared enterprise between the teacher and student has always been, and always will be, fundamental to effective pedagogic practice.



As such, we have three key recommendations:

1. *For policy developers and professional development providers:* It is necessary for teachers to be supported with professional development that emphasises the philosophical underpinnings of effective practice, in particular the emphasis on engaging students as owners of their learning and to purposefully link these practices to policy. Fenwick and Cooper (2012) state:

current standards-based reforms in Australia will do little to improve student achievement unless there is a national commitment to the provision of professional learning that presents the ideas and research underpinning the changes and aims to challenge some current pedagogies that involve matching curriculum and expectations with perceptions of students' current capacities (p. 359).

Policy and practice need to connect to the conceptual understandings informed by research rather than reducing the spirit of AfL to a scattered collection of strategies.

2. *For professional development providers, school leaders and researchers:* Opportunities for facilitated professional dialogue need to be incorporated as an aspect of professional practice as teachers develop shared understandings of the policy context in which they work. A report from the OECD includes informal teacher dialogue as one of the practices that has the greatest impact on teaching practice (Organisation for Economic Co-operation and Development (OECD) 2011).
3. *For teachers and school leaders:* Collaboratively annotating student samples with reference to the achievement standards needs to be included in the activities that occur during the planning process at the beginning of a semester. This will involve time for focussed and purposeful dialogue on using assessment to inform teaching and learning. This is necessary for teachers to develop a shared understanding of the standard that will inform the planning of discrete curriculum units of work and the flow on to teaching practice. Teachers' clear understanding of the qualities that provide evidence of a standard when shared with their students will empower students to also come to know the teacher's interpretation of the standard and use this knowledge to inform their self-assessment and achieve their learning goals.

## References

- Absolum, M., Flockton, L., Hattie, J., Hipkins, R., & Reid, I. (2009). *Directions for assessment in New Zealand (DANZ): Developing students' assessment capabilities*. Wellington: Ministry of Education. <http://assessment.tki.org.nz/Media/Files/Directions-for-Assessment-in-New-Zealand>. Accessed 21 Feb 2015.
- Adie, L., & Willis, J. (2014). Using annotations to inform an understanding of achievement standards. *Assessment Matters*, 6, 112–136.
- Assessment Reform Group. (2002). *Testing, motivation and learning*. Cambridge: The Assessment Reform Group.

- Australian Curriculum, Assessment and Reporting Authority. (2013a). *NAPLAN*. <http://www.nap.edu.au/naplan/naplan.html>. Accessed 21 Feb 2015.
- Australian Curriculum, Assessment and Reporting Authority. (2013b). *NAPLAN Infographic*. [http://www.nap.edu.au/verve/\\_resources/Acara\\_NAPLAN\\_Infographic%28V4-2%29.pdf](http://www.nap.edu.au/verve/_resources/Acara_NAPLAN_Infographic%28V4-2%29.pdf). Accessed 21 Feb 2015.
- Australian Curriculum, Assessment and Reporting Authority. (2014). *The Australian curriculum*. <http://www.australiancurriculum.edu.au/>. Accessed 21 Feb 2015.
- Australian Curriculum, Assessment and Reporting Authority. (2015). *Implications for teaching, assessing and reporting*. <http://www.australiancurriculum.edu.au/overview/implications-for-teaching-assessing-and-reporting>. Accessed from multiple curriculum sites 02 Nov 2015.
- Australian Institute for Teaching and School Leadership. (2014). *Australian professional standards for teachers*. <http://www.aitsl.edu.au/australian-professional-standards-for-teachers/standards/overview/organisation-of-the-standards>. Accessed 21 Feb 2015.
- Ball, S., Maguire, M., Braun, A., & Hoskins, K. (2011). Policy subjects and policy actors in schools: Some necessary but insufficient analyses. *Discourse*, 32(4), 611–624.
- Berliner, D. (2010). *High-stakes assessments and the narrowing of the outcomes of schooling and students' minds*. Keynote paper presented at the blind assessor: Are we constraining or enriching student learning? Symposium, Sydney NSW. [http://sydney.edu.au/education\\_social\\_work/professional\\_learning/resources/2010\\_papers.shtml#sthash.91JoTqco.dpuf](http://sydney.edu.au/education_social_work/professional_learning/resources/2010_papers.shtml#sthash.91JoTqco.dpuf). Accessed from multiple curriculum sites 02 Nov 2015.
- Bernstein, B. (2003). *Class, codes and control: The structuring of pedagogic discourse* (Vol. 4). London: Routledge.
- Black, P. (2007). Full marks for feedback. In *Make the Grade* (pp. 18–21). Spring edition.
- Earl, L., & Timperley, H. (2014). Challenging conceptions of Assessment. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 325–336). Dordrecht: Springer.
- Education Scotland. (2015). *Principals of assessment*. <http://www.educationscotland.gov.uk/learningteachingandassessment/assessment/about/principles/learnerengagement.asp>. Accessed 21 Feb 2015.
- Fenwick, L., & Cooper, M. (2012). Prevailing pedagogies for classes in low SES contexts and the implications for standards-based reform in Australia. *Australian Educational Research*, 39, 349–361.
- Fitzgerald, T., & Savage, J. (2013). Scripting, ritualising and performing leadership: Interrogating recent policy developments in Australia. *Journal of Educational Administration and History*, 45(2), 126–143. doi:10.1080/00220620.2013.769432
- Fleming, J., & Kleinhenz, E. (2007). *Towards a moving school: Developing a professional learning and performance culture*. Camberwell, Australia: ACER Press.
- Freebody, P. (2005). *Background, rationale and specification: Queensland curriculum, assessment and reporting framework*. Queensland, Australia: Queensland Government, Department of Education and the Arts.
- Goldstein, H. (2010, January 27). UK academic warns against Australian school league tables (Audio podcast). [http://blogs.abc.net.au/canberra/2010/01/uk-academic-warns-against-school-league-tables.html?site=canberra&program=\(none\)](http://blogs.abc.net.au/canberra/2010/01/uk-academic-warns-against-school-league-tables.html?site=canberra&program=(none)). Accessed 03 Nov 2015.
- Griffin, P. (2007). The comfort of competence and the uncertainty of assessment. *Studies in Educational Evaluation*, 33(1), 87–99.
- Harlen, W. (2007). *Assessment of learning*. London: Sage.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Oxon: Routledge.
- Klenowski, V. (2009). Assessment for learning revisited: An Asia-Pacific perspective. *Assessment in Education: Principles, Policy and Practice*, 16(3), 263–268.
- Lingard, B., Mills, M., & Hayes, D. (2006). Enabling and aligning assessment for learning: Some research and policy lessons from Queensland. *International Studies in Sociology of Education*, 16(2), 83–103.

- Looney, A. (2014). Assessment and the reform of education systems. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 233–248). Dordrecht: Springer.
- Luke, A. (2011). Generalizing across borders: Policy and the limits of educational science. *Educational Researcher*, 40(8), 367–377.
- Marshall, B., & Drummond, M. (2006). How teachers engage with assessment for learning: lessons from the classroom. *Research Papers in Education*, 21(2), 133–149.
- Marzano, R. J. (2007). *The art and science of teaching: A comprehensive framework for effective instruction*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Masters, G. (2009). *A shared challenge: Improving literacy, numeracy and science learning in Queensland primary schools*. Australian Council for Educational Research. <http://education.qld.gov.au/mastersreview/pdfs/final-report-masters.pdf>. Accessed 21 Feb 2015.
- Matters, G., & Masters, G. (2014). *Redesigning the secondary-tertiary interface: Queensland review of senior assessment and tertiary entrance*. Report submitted to the Queensland Minister for Education, Training and Employment, Sept 2014. <http://www.acer.edu.au/queensland-review>. Accessed 21 Feb 2015.
- New Zealand Ministry of Education. (2011). *Ministry of education position paper: Assessment (schooling sector)*. Wellington: Learning Media Ltd. [http://www.minedu.govt.nz/~/media/MinEdu/Files/TheMinistry/MOEAssessmentPositionPaper\\_October11.pdf](http://www.minedu.govt.nz/~/media/MinEdu/Files/TheMinistry/MOEAssessmentPositionPaper_October11.pdf). Accessed 21 Feb 2015.
- Organisation for Economic Co-operation and Development (OECD). (2011). *Building a high quality teaching profession: Lessons from around the world*. Background report for the International Summit on the Teaching Profession: New York 2011. OECD publishing, Paris. <http://www2.ed.gov/about/inits/ed/internationale/background.pdf>. Accessed 21 Feb 2015.
- Popham, W. J. (2014). Looking at assessment through learning-colored lenses. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 183–194). Dordrecht: Springer.
- Pryor, B., & Crossouard, B. (2008). A socio-cultural theorisation of formative assessment. *Oxford Review of Education*, 34(1), 1–20. doi:10.1080/03054980701476386
- Queensland Curriculum & Assessment Authority. (2014). *School-based assessment. The Queensland system*. Queensland Government. [https://www.qcaa.qld.edu.au/downloads/approach/school-based\\_assess\\_qld\\_sys.pdf](https://www.qcaa.qld.edu.au/downloads/approach/school-based_assess_qld_sys.pdf). Accessed 21 Feb 2015.
- Queensland Curriculum and Assessment Authority. (2015). *Queensland comparable assessment tasks (QCATs)*. Queensland Government. <https://www.qcaa.qld.edu.au/3163.html>. Accessed 02 Nov 2015.
- Queensland Studies Authority. (n.d.) *A new perspective: Assessment for learning*. [https://www.qcaa.qld.edu.au/downloads/p\\_10/as\\_afl\\_new\\_perspective.docx](https://www.qcaa.qld.edu.au/downloads/p_10/as_afl_new_perspective.docx). Accessed 21 Feb 2015.
- Sadler, R. (1998). *Letting students into the secret: Further steps in making criteria and standards work to improve learning*. Paper presented at the Annual Conference for State Review Panels and District Panel Chairs, Brisbane. [http://www.qsa.qld.edu.au/yrs11\\_12/assessment/docs/criteria.pdf](http://www.qsa.qld.edu.au/yrs11_12/assessment/docs/criteria.pdf). Accessed 21 Feb 2015.
- Singh, P. (2002). Pedagogising knowledge: Bernstein's theory of the pedagogic device. *British Journal of Sociology of Education*, 23(4), 571–582.
- Stiggins, R. (2002). Assessment crisis: The absence of assessment for learning. In *Phi Delta Kappan* (pp. 758–765). June.
- Swaffield, S. (2011). Getting to the heart of authentic Assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 18(4), 433–449. doi:10.1080/0969594X.2011.582838
- Tuinamuana, K. (2011). Teacher Professional standards, accountability, and ideology: Alternative discourses. *Australian Journal of Teacher Education*, 36(12), 72–82.
- Willis, J. (2011). Affiliation, autonomy and assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 18(4), 399–415. doi:10.1080/0969594X.2011.604305
- Willis, J., & Adie, L. (2013). Negotiating the meaning of achievement standards in the Australian curriculum. *Curriculum Perspectives*, 33(1), 52–62.

- Willis, J., & Adie, L. (2014). Teachers using annotations to engage students in assessment conversations: Recontextualising knowledge. *Curriculum Journal*, 25(4), 495–515. doi:[10.1080/09585176.2014.968599](https://doi.org/10.1080/09585176.2014.968599)
- Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *Australian Educational Researcher*, 40(2), 241–256. doi:[10.1007/s13384-013-0089-9](https://doi.org/10.1007/s13384-013-0089-9)

# Chapter 4

## Effective Enactment of Assessment for Learning and Student Diversity in Australia

J. Joy Cumming and Fabienne M. Van der Kleij

**Abstract** This chapter examines implementation of Assessment for Learning (AfL) for diverse students, including students from diverse language and cultural backgrounds, and, in particular, for students with disabilities that affect their learning. Australian national education policy endorses AfL as effective teacher assessment practice. Australian education policy also promotes educational equity for all students, regardless of linguistic or cultural background, or disability, not only in terms of access to schooling but also in terms of access to a high quality and challenging education. This chapter provides an overview of Australian equity policy, the Australian federated system of education policy development and responsibilities, and the recent policy initiative of a national curriculum as context for AfL practice. We identify core principles of AfL with respect to teacher–student interactions for consideration of issues for curriculum and AfL implementation for diverse students. We then provide an overview of international research on AfL for these students. Four Australian examples of pedagogical interactions between teachers and students with disabilities are examined in terms of equity, curriculum, and AfL policy expectations. We conclude that to enact effective AfL for these students, while policy can and should provide an enabling framework, much at present depends on the individual teacher’s in-depth knowledge of students. To achieve effective and equitable implementation of AfL, further resources and professional development support are needed. Suggested guidelines are provided to enhance policy, practice, and research in AfL for diverse students.

### 4.1 Introduction

Australia has joined the international community in identifying Assessment for Learning (AfL) as a significant policy pillar in the focus on improvement of student learning, influenced, at least in part, by the early works of Crooks (1988), Sadler

---

J.J. Cumming (✉) · F.M. Van der Kleij  
Australian Catholic University, Brisbane, Australia  
e-mail: joy.cumming@acu.edu.au

(1989), Black and Wiliam (1998), and the Assessment Reform Group (ARG) (2002). The major policy statement of Australia's national education goals, the *Melbourne Declaration on Educational Goals for Young Australians* (Ministerial Council on Education, Employment, Training and Youth Affairs [MCEETYA] 2008), commits to 'world-class' assessment (p. 10). This incorporates external assessments but also classroom-based teacher assessment practices that focus on:

- assessment for learning—enabling teachers to use information about student progress to inform their teaching
- assessment as learning—enabling students to reflect on and monitor their own progress to inform their future learning goals
- assessment of learning—assisting teachers to use evidence of student learning to assess student achievement against goals and standards. (p. 14)

While assessment *of* learning addresses teachers' collection and use of assessment evidence for summative assessment purposes, Australia's national education goals explicitly promote both assessment *for* and *as* learning, resonating with international understandings of AfL.

The ARG (2002) defined AfL as 'the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there' (p. 2).

AfL has become intertwined in literature and practice with research on formative assessment. In this chapter, our AfL focus is on policies that promote ongoing teacher–student pedagogical interactions and feedback loops intended to improve individual student learning outcomes—the classroom dialogue both verbal and visual. Hayward and Spencer (2014) identify '*dialogue* (pupil–teacher and pupil–pupil)' as the 'key recurring element' to independent learning assessment processes that challenge 'learners to reflect on their own thinking and to make unconscious learning processes overt, so that they can be considered, discussed and improved' (p. 17).

The ARG identified four AfL principles that establish a framework for these teacher–student interactions and dialogues:

- focus on how students learn
- promotion of understanding of goals and criteria
- guidance on how to improve
- development of capacity for self-assessment (ARG 2002).

Guidance on how to improve, including feedback, is critical. Drawing on their own earlier work (Black and Wiliam 1998), as well as work by Ramaprasad (1983), Sadler (1989) and Black and Wiliam (2009) identified three key processes in classroom feedback to '[move] learners forward'—the need to establish where 'learners are in their learning,' 'where they are going,' and 'what needs to be done to get them there' (pp. 7–8). Black and Wiliam (1998) note Perrenoud's perception that these processes require 'an incursion into the representation and thought

processes of the pupil, to accelerate a breakthrough in understanding, a new point of view or the shaping of a notion which can immediately become operative' (Perrenoud 1998, p. 97). These interactions are necessarily shaped by the identity and context of the learner. Immediately, it is clear that these basic, yet conceptual, understandings of processes that improve learning may present challenges for teaching diverse students.

In this chapter, we identify five elements which engage all these principles and processes (based on ARG 1999; William and Thompson 2007) to consider their implementation for diverse students, that is, students with linguistic and cultural difference, and with disabilities, and the extent to which AfL practice with these students can achieve Australian assessment policy goals. The five elements are:

- determining how students learn and learning progressions
- sharing learning expectations
- questioning to gauge knowledge and understanding
- provision of feedback for learning improvement
- development of student capacity for self-assessment.

The following sections (i) describe the Australian education policy context for equity, curriculum, and AfL; (ii) provide an overview of available research on AfL for diverse students; (iii) consider implementation of the five identified AfL elements for students with disabilities through four examples; (iv) discuss implications for diverse students and provide recommendations for policy, practice, and research.

## 4.2 Equity, Curriculum, and Assessment Policy Contexts for Australian Schools

Australian school students have diverse cultural, social, and linguistic backgrounds. For example, nearly 30 % of students in New South Wales' government schools have English as a second language (Department of Education and Communities 2011). Over five per cent of students identify as Indigenous Australians of Aboriginal or Torres Strait Islander culture (Australian Bureau of Statistics 2015), many of whom may also speak a home language other than English. More than one in ten Australian students may have identified disabilities (PricewaterhouseCoopers 2013), the majority attending mainstream schools with varying degrees of support. Australian students progress through school with their age (social) cohort, not according to their achievement level. Hence, every Australian school classroom will present challenging complexity of diversity and individual student needs for teachers.

The Melbourne Declaration commits to equity and excellence in education for all, not only in discipline learning needs but also in affective outcomes that impact on future life quality (MCEETYA 2008). Students are to be successful learners, confident and creative individuals, and active and informed citizens. The

Declaration identifies that Indigenous students, students from disadvantaged socioeconomic backgrounds, and students with disability are not achieving equitable outcomes. Strategic initiatives to improve equity in opportunity and outcomes for these students are highlighted in the Declaration as priority policy initiatives for Australian education.

These policy initiatives are supported by Australian legislation. Antidiscrimination legislation exists at federal, state, and territory levels to prevent discrimination on grounds of characteristics such as race, culture, gender, sexuality, disability, or religion. Specific subordinate antidiscrimination legislation, the *Disability Standards for Education 2005* (Attorney-General 2005), addresses education provision for students with disabilities. Overall, these students are to be provided with access to curriculum and programs on a similar basis to students without disabilities, and more specifically, are to be assessed in ways appropriate to the circumstances that enable them to demonstrate their learning.

The Australian policy framework, supported by legislation, therefore mandates and endorses equitable and high quality world-class education for all students. It prioritises AfL as world-class assessment practice to benefit learning of all students. As a corollary, AfL is therefore an implicitly critical assessment practice to enable students from diverse language and cultural backgrounds or with disability to achieve equitable learning outcomes.

#### ***4.2.1 The Australian Curriculum and Student Diversity***

Australia has a federated system of education responsibility. The *Melbourne Declaration on Educational Goals for Young Australians* (MCEETYA 2008) is the third and most recent national statement of education goals since the first in 1989. Their significance for Australian education is the collaboration of all state and territory ministers and the federal minister of education to develop a common education policy framework for all Australian students.

The Declaration's goal of a common curriculum framework has been achieved; a new national Australian Curriculum has been under development by the Australian Curriculum, Assessment and Reporting Authority (ACARA) since 2008. This Curriculum has a three dimensional structure: learning areas (disciplines/subjects); General Capabilities (essential 21st-century skills); and cross-curriculum priorities. The curriculum is standards-based and provides qualitative descriptors of expected achievement standards for each Year level in each learning area.<sup>1</sup> While the Australian Curriculum provides the common content framework, curriculum implementation and school assessment remain the responsibility of state and territory authorities.

---

<sup>1</sup>The Australian Curriculum along with supporting resources and guidance are accessible online ([www.australiancurriculum.edu.au](http://www.australiancurriculum.edu.au)).



The Australian Curriculum addresses the Melbourne Declaration goal of equity and student diversity. ACARA (2013b) is ‘committed to development of a high-quality curriculum for all Australian students that promotes excellence and equity in education. All students are entitled to rigorous, relevant and engaging learning programs drawn from challenging curriculum that addresses their individual learning needs’ (p. 4).

A core Curriculum proposition is that ‘each student can learn and... the needs of every student are important’ (ACARA 2012, p. 10). Guidelines suggest application of the following process to address students’ diverse learning needs:

1. Identify suitable learning area content considering the students’ age.
2. Modify teaching if needed drawing from different Year level content, using the General Capabilities and/or cross-curriculum priorities to modify the learning focus, or align individual learning goals with age-appropriate learning content.
3. Assess students against curriculum standards or according to individual learning goals (ACARA 2013b).

Guidelines and a language learning continuum are available for students with English as an Additional Language or Dialect, with advice that these students ‘may require additional time and support’ to learn (ACARA 2015). Literacy and Numeracy continua have also been developed within the General Capabilities to assist individual planning for students with disabilities. The starting proposition, however, is that individualised curricula for students should be based on age-appropriate content, that is, while student learning may not be at the same level as that of other students, students will still experience the full curriculum. Although ACARA does not have responsibility for implementation of school assessment, a state and territory responsibility, the Australian Curriculum website provides exemplars for practical guidance in such adaptive planning. State and territory authorities and school sectors (Government, Independent [both faith and nonfaith based] and Catholic) are expected to provide teachers with more specific policy and support.

While the Melbourne Declaration indicates that teachers should engage with assessment *for* and *as* learning, the Australian Curriculum states only that teachers should practise ongoing formative classroom assessment to continuously inform teaching and learning. Again, specific guidance is a state and territory responsibility.

We note that Australia may present a different teacher practice scenario from other countries. Given historical development of curriculum and policy at state and territory systemic levels and legislative and financial controls, school practice must be compliant with these curriculum and policy expectations (Cumming and Mawdsley 2012). While enactment of curriculum and policy at the classroom level will always differ from official expectations, they will be aligned. A study already undertaken of teacher implementation of the Australian Curriculum in the subject English found that while teachers developed their own resources, approximately three-quarters of the study’s respondents identified state and territory curriculum documents as important for long-term planning, with half using them in short-term planning (Albright et al. 2013).

### ***4.2.2 Assessment for Learning Policy and Guidance at State and Territory Level for Diverse Students***

Compatible with national policy goals, assessment policies of each Australian state and territory education authority endorse AfL as a key component of expected teacher assessment practice. Australian state and territory authority websites refer to the need for all teachers to undertake assessment for, as, and of learning, frequently listing the Black and Wiliam (1998) or ARG (2002) generic principles. For example, the Queensland Curriculum and Assessment Authority references AfL principles within all curriculum areas. The Board of Studies of New South Wales identifies AfL as ‘quality assessment that has had world-wide success in enhancing teaching and improving student learning’ (2015, paragraph 1). The Department of Education and Early Childhood Development in Victoria identifies the primary purpose of assessment as improvement of student learning (2013). Online professional development modules addressing generic AfL principles are provided to guide in-school workshops.

State and territory education authorities also provide policy and guidelines addressing equity and inclusive practices in education and provision for diverse students. General principles are that instruction should meet each student’s needs, with suggestions provided for different forms of summative assessment evidence, such as observations and anecdotal observations, or modifications or adjustments in formal summative assessments. What is missing at the state and territory level are the policies and guidance for teachers that integrate AfL with the learning characteristics and needs of these students, and provide specific consideration of implications of the interaction of AfL and the learning of these students. Thus, the only practical guidance available to Australian teachers for working with students from diverse language and cultural backgrounds or with disability and AfL is at the national level.

The following section considers research on AfL for these students to inform what policy guidance should be provided for the Australian national, state, and territory educational context.

## **4.3 Research on AfL for Diverse Students**

Limited research has addressed the role of AfL for the diversity of students encountered in classrooms today, including students with different language backgrounds from the language of instruction, different cultural backgrounds, or students with disabilities. Black and Wiliam’s informing review of empirical research (1998) reported positive effects for the majority of students, including low attaining students. However, the learning outcomes and diversity of students, beyond characterisation as disadvantaged or low achieving, investigated in the review are limited. Fuchs and Fuchs (1986) reported positive effects for

‘systematic’ formative assessment for students with disabilities (predominantly ‘mildly handicapped’) operationalised as twice-weekly ‘curriculum-based data collection [excluding non-academic behaviours] ... with decisions concerning the adequacy of programs formulated on an individual, not a group, basis’ (p. 201). Fuchs et al. (1997) found that task-focused goals and self-referenced assessment feedback based on a weekly classroom test improved outcomes for low achieving students but not for students with identified learning disabilities. While these studies incorporated aspects of AfL, they did not reflect ongoing daily teacher–student learning interactions.

A methodological issue is that quantitative empirical research on effective practices for diverse students, especially children with disabilities, generally treats students with diverse characteristics as a homogenous group, obscuring positive individual outcomes (Cumming 2012; Pitoniak and Royer 2001). The nature of knowledge construction (ways of knowing), different learning progressions, interaction of assessment and culture, and the individuality and idiosyncrasy of learners with disabilities present challenges for effective AfL because of the different ways in which diverse learners both learn and are able to demonstrate their learning (Abedi 2010; Bourke et al. 2011; Cumming 2012; Ravet 2013).

Trumbull and Lash (2013) provide a comprehensive overview of formative assessment principles and potential unintentional effects for diverse students. They identify formative assessment as a process aligned with teaching designed to examine the nature of students’ understanding and to advance student agency. As Trumbull and Lash note, emphasis on ‘closing the gap’ through feedback necessarily requires implicit or explicit conceptualisation of the nature of the intended learning progression. However, they also note that few curriculum sequences are empirically developed, although in Australia, as elsewhere, such sequences underpin most instructional planning and assessment. Diverse students, such as students with cultural diversity or learning disabilities, may not fit with standard expected learning progressions. Baird et al. (2014) noted that alignment of ‘current understandings of formative assessment’ with different learning theories would ‘be reflected in differing formulations and practices’ (p. 30). This must similarly apply for different progressions of learning and *different* students.

### 4.3.1 AfL, Language, and Cultural Diversity

Trumbull and Lash (2013) examine implications of formative assessment principles for students with language backgrounds different from the language of instruction and with different cultural knowledge and experience. To be successful learners, students must develop several knowledge structures simultaneously, the instructional language, the intended curriculum content, and, often, the culture of interaction. Complex linguistic text may pose a barrier to demonstration of knowledge by students learning in a second language (Abedi 2010). Using different communication modes such as nonlinguistic or visual modes could provide more valid

information on which to base feedback. Thus, the key recurring element of dialogue in AfL may be problematic in these learning contexts.

Differences in the home *cultural script* of parent-child conversations may also mean students need scaffolding in the classroom discourse promoted in generic discussions of AfL processes. Students from different cultural environments may have different understandings of social roles within classrooms. Teacher questioning intended to elicit student understanding can be impacted by cultural differences, leading teachers to wrongful interpretation of the extent of student learning. As an example, in some cultures, direct questioning or being singled out are not culturally appropriate. Trumbull and Lash (2013) note that among Native American groups, dichotomous *right-wrong* approaches are not cultural practice: ‘Teachers not privy to the communication norms in some communities may at times be introducing non-target [learning goals] into assessment by using ... formative assessment practices that are most accepted (e.g., questioning students during a whole group discussion)’ (p. 12).

Cultural reactions to praise may be another source of difference. Hence, teachers need to know their students as individuals and tailor practices to be culturally sensitive and appropriate. No research has been identified on AfL and Australian Indigenous students. In one study of teacher assessment that interviewed over 100 teachers of these students (Klenowski 2014), some reported use of approaches identified as AfL practice, for example, asking open-ended questions, providing feedback to inform students how to improve, and student agency in assessment processes. Cultural interactions, however, were not the focus of the study.

### 4.3.2 *AfL and Students with Disabilities*

The Melbourne Declaration (MCEETYA 2008) addresses equitable educational opportunity for all students, including students with disability, and the ideal of AfL, but without explicitly linking the two. The Australian Curriculum implicitly links the two but without a clear policy bridge. The European Agency for Development in Special Needs Education (EADSNE) (2009) specifically examined AfL for students with special education needs and noted that foundational AfL work was based on students without such needs. Drawing on conversations with project experts, EADSNE noted the importance of AfL principles for all students, our own starting point. The experts considered that classroom interactions such as questioning were possible with students with special education needs as long as the question and answer modes used different stimuli, for example ‘visual versus verbal stimuli’ (p. 5), and modes compatible with students’ capabilities. Encouraging self-reflection and development of metacognitive skills were identified to be as desirable for students with special needs as for other students. A core practical issue identified for AfL, but unresolved, is how to involve learners with ‘profound difficulties’ (p. 4) in feedback loops and self-reflection.

Ravet (2013) has undertaken a comprehensive critical evaluation of the suitability of recommended generic AfL principles for students on the autism spectrum, drawing mainly on Black and Wiliam's work (e.g., 1998, 2009). She identified that many principles, including 'opportunities for pupil interaction, teacher/pupil dialogue, high quality feedback to pupils and pupil self and peer assessment' (p. 950), were problematic and most likely counter to improving these students' learning. Students with autism have different ways of knowing and interacting—'a different way of perceiving information; a different way of experiencing the world, a different way of coding, storing and retrieving in memory; and a different role of the emotions in processing these elements' (p. 953, citing Powell and Jordan 2012). Common characteristics of autism include difficulties in communications and interactions with others, behavioural concerns, and tendency to focus on detail rather than the whole. However, students can also have learning strengths to draw on, including deep knowledge, objectivity, enjoyment of individual work, strong focus, and persistence.

Ravet (2013) concludes that the 'majoritarian' view of learning (p. 954) promoted through generic approaches results in issues related to (1) inferences from evidence of student learning and (2) communication. Firstly, knowing where students are in their learning is integral to AfL, but limited teacher understanding of autism may lead to biased inferences that the teacher is not aware of, with negative impact on validity of subsequent instructional actions. The inferential process in day-to-day classroom judgements of learning is highly influenced by teachers' subconscious beliefs and intuition (Bennett 2011; Ravet 2013). Secondly, communication difficulties for many learners with autism pose challenges for interactive classroom processes such as peer assessment, social interaction, and metacognitive feedback (Ravet 2013). Ravet proposes adaptations to Black and Wiliam's (1998) formative assessment activities to be inclusive for students on the autism spectrum. These adaptations in practice require teachers both to be flexible and to have thorough understanding of autism.

While feedback is a key process in AfL, in practice, feedback often takes the form of praise, identified in research as least effective for student learning. Feedback is identified as most effective when it focuses on the learning (the task, process, or metacognitive strategy), not the student (Black and Wiliam 1998; Hattie and Timperley 2007). Research has established that the effectiveness of feedback for learning relates to characteristics of the feedback (content, timing), the types of learning outcomes (Shute 2008; Van der Kleij et al. 2015), and learner characteristics such as ability levels (Shute 2008). However, this research, again, has focused on what is effective for the majority of learners. The further question, then, is what type of feedback is most effective for students with different nature and extent of disability.

A further issue raised in education of students with disabilities is the construction of Individual Education Programs (IEPs). Many Australian researchers have criticised these in the past as narrowing curriculum opportunities and creating a *deficit* approach to learning, hindering inclusion, and encouraging low expectations by teachers (Carrington and MacArthur 2012; Shaddock et al. 2007). The Melbourne

Declaration and Australian Curriculum directly address this issue through principles of high expectations for all. Despite good intentions, however, individualised plans may result in exclusion of students from potential classroom interactions with peers and AfL opportunities—the ‘IEP goals become a separate curriculum’ (Carrington and MacArthur 2012, p. 278).

#### 4.4 AfL for Students with Disabilities: Learnings from Australian Examples

In this section we draw on four examples of pedagogic interactions between teachers and students with disabilities to discuss their consequences for AfL implementation.

The first two examples draw upon Australian classroom research data. The following two examples are drawn from the national resources available to Australian teachers to support enactment of equity policy in the Australian Curriculum, *Illustrations of Personal Learning*, to guide teachers in personalising curriculum for students with diverse needs and to ‘demonstrate how the integrity of the learning areas ...[can be] maintained while addressing individual learning needs’ (ACARA 2013b, p. 18).

While none of the examples discussed in this section were intentionally based within AfL practices, each raises consideration of at least one of the five core elements on which we have focused:

- determining how students learn and learning progressions
- sharing learning expectations
- questioning
- provision of feedback for learning improvement
- development of student capacity for self-assessment.

*The first example* involves an inclusive classroom learning activity incorporating a performance sheet with criteria (dimensions of performance) and standards (dimensions of quality) to establish parameters of a writing task. The teacher is implementing assessment that can be used as assessment *of* learning within a framework that uses AfL practices to scaffold student learning, in accord with general state (Queensland) assessment policy expectations. The student, in the last year of primary school, Year 7 (approximately 12–13 years old), had ongoing literacy learning needs (Colbert and Cumming 2014). This teacher–student interaction sequence draws on a longitudinal project reported elsewhere (Wyatt-Smith et al. 2007). It involves interviews with and between the teacher and student, and classroom work collected over the year. Although the student engaged with the same content and completed the same task as other students, even prior to current Australian Curriculum expectations, the teacher modified his performance expectations by using simplified and fewer criteria. The student was assessed against

modified standard categories: *developing*, *developed*, and *highly developed*. The performance sheet formed the basis of verbal and visual dialogue between the teacher and student on several occasions, scaffolding the student's learning and work, with expectations refined iteratively over time. Both the teacher and student provided assessments of his progress, by shading or dots on the sheet, engaging student reflection and agency. The student identified specific improvement goals based on the performance expectations. As time progressed, he asked for expectations to be raised both in terms of the number of criteria and standards of performance, and made more similar to those of his peers.

While the available data did not provide information on teacher–student questioning or verbal feedback on the task, this example demonstrates AfL principles for a student with a literacy learning disability within mainstream curriculum. The performance sheet detailed expectations, was used to provide ongoing feedback, enabled adaptation of goals as the student's learning progressed, was used to scaffold the classroom dialogue, and the student's own motivation, and served to improve his engagement and facilitate his learning. The assessment processes were in accord with curriculum and policy expectations for both a student with specific learning needs and AfL. However, the critical element was the teacher's knowledge of the student's capabilities and flexibility to adapt expectations to match these initially, shaping feedback within an appropriate learning progression. Critical also was the student's willingness to engage with feedback and improve his learning. While written literacy was a difficulty for this student, verbal interactions were not.

*The second example* is taken from a small research project undertaken in 2014 investigating the usability of a new State policy development, a curriculum document, the Guideline for Individual Learning (GIL) (Queensland Curriculum and Assessment Authority [QCAA] 2014) to implement the Australian Curriculum policy expectations for students with disability and frame their learning. The GIL is designed to align education goals for students with mild, moderate to severe intellectual and/or physical disabilities with the Australian Curriculum. Students completing Years 11 and 12, the last two years of secondary schooling (approximately 17–19 years old), receive a Queensland Certificate of Individual Achievement (QCIA) on successful attainment of their individual goals. The GIL reflects the core policy proposition of the Australian Curriculum that the starting point for individual student curricula is age-appropriate content. Assessment advice within the GIL provides an explicit, but not elaborated, link between generic formative assessment principles and learning of students with disabilities; it is stated that assessment should 'promote, assist and improve teaching and learning' by providing 'regular feedback to students about how they can improve their learning' (p. 8).

Semi-structured interviews with principals and teachers in three schools (two special education and one inclusive education setting) discussed the penultimate version of the GIL and its implications for assessment. This example involves three students in one of the schools: a student with a speech language disability, a student with an intellectual impairment, and a student with autism spectrum disorder (high anxiety). The three students were undertaking studies towards a vocational

certificate as well as the QCIA in an inclusive education setting. Summative assessment and reporting for each of these students occurred through judgements against qualitative criteria and standards as well as a competency checklist. However, a number of areas directly relating to AfL principles emerged from the interviews.

In this setting, teachers worked together with the GIL framework to translate coarse-grained Australian Curriculum standards into curriculum-consistent but different achievable learning goals to suit each student. They tracked the learning gains of the students, and reviewed goals and lesson plans every few weeks. Continuous feedback was provided verbally (as reported by staff) and through students' work. Goals from the overall curriculum were chunked into smaller manageable 'bites' that were scaffolded until evidence of successful achievement could be documented. This was seen as critical to student success. Scaffolding related not only to development of specific vocational skills but also to students' development to independent learning. For the student with autism, this was undertaken first through the student observing her teacher undertaking a task, then the teacher and student working jointly, until the student had confidence and skill to work independently. AfL emphasises explicit goal setting and sharing of criteria for success. Often in mainstream curriculum, as in the Australian Curriculum, final summative performance expectations identify large-step learning goals for the end of a school semester but do not elaborate implicit or explicit stages to be achieved on route to these goals (Popham 2008). In this example, the shared bite-sized learning goals reflected simultaneously the underpinning curriculum, that is, the content or skill that was the overall goal, and the learning progressions for each student to develop independent work skills. The new policy development of the GIL with explicit curriculum, generic advice on assessment, and use of rubrics to judge student achievement enabled this work by teachers. However, once more their effective enactment of the policy was still very dependent on their own experience and strategies in working with these students.

Self-assessment is a key principle of AfL, endorsed in assessment policies across Australian education with very limited empirical investigation for students with disabilities. A further observation from this example was that student self-assessment capability, becoming 'reflective and self-managing' (ARG 2002, p. 20), was achievable for these three students. Digital portfolios, recorded on iPads, provided the achievement evidence base. Teachers shared with students the outcomes they were to achieve and gave students responsibility to determine their own evidence. Students took photos of their work, screenshots of internet search histories, or were videoed undertaking tasks such as reading a newspaper aloud. Students emailed these to the staff member compiling the portfolios. Given the multitude of learning evidence collected, the students were also charged by teachers with determining what evidence best represented their learning outcomes. Staff reported that not only was better evidence collected than previously when undertaken by classroom aides, student agency in data collection greatly increased their engagement with their learning. The use of technology proved to be highly effective for these students with disabilities.



I've never seen kids who have this kind of impairment interact with technology so well... all because we're saying we need this kind of evidence and they're aware of how technology can provide that evidence. (Head of Learning–Learning Enhancement)

As noted, the Australian Curriculum is intended to provide a curriculum for all students, with several policy guidelines available online to assist teachers in adapting the Curriculum to suit individual learning needs for diverse students. The final two examples are drawn from online vignettes provided as Australian Curriculum teacher resources on identification and implementation of appropriate learning pathways for students with moderate to severe intellectual and physical disabilities.

*Vignette 1* (ACARA 2013a) explores one teacher's curriculum planning for seven students with moderate to significant intellectual disability, ranging from five to nearly 13 years old. All students have individualised curriculum plans, aligned with Australian Curriculum General Capabilities continua. Communication goals are a strong focus. In the vignette, the teacher combines age-appropriate curriculum content with learning goals suited to her students' individual education programs, aligning her judgement of student learning to these multiple individual goals. Several tools support communication in teacher–student interactions, including symbols, iPad applications, and a communication book. Physical objects are used for effective questioning. For example, students are asked 'what's your favourite boat or ship in the book?' [book with pictures of boats]; a child makes a sound, the teacher follows up:

With your pointing finger [teacher touching child's finger].  
That's right, with your pointing finger.  
Which boat was your...  
(GASPS) You're touching that one which is the...

The teacher's feedback is on both content and reinforcement of the process for communication. Student agency in the learning process involved freedom to choose a writing topic related to the boat theme. In the second part of the vignette, the teacher conducts a science experiment with these students, focusing on forces and two floating boats, contextualised through a recent boating experience by the students. Variables such as the types of sail and weights in the boats were changed, and students were asked to predict which boat would finish first. Again, student questioning takes place in the form of offering them a concrete choice: a blue or a red boat. While questions are phrased directly, the teacher verbalises fully the differences between the boats, encompassing sophisticated scientific principles. Through feedback the teacher scaffolds students' behaviour and communication, as well as their cognitive development in line with the curriculum. She has clear understanding of the learning progressions of these students in all these areas. Given communication constraints, while the teacher shared learning expectations with the students, she must infer their understanding and internal processes through her knowledge of the students (Perrenoud 1998). Her questioning and feedback are constant, using verbal, visual, and physical stimuli (EADSNE 2009). Student agency is facilitated, not necessarily through self-assessment, but through student choice in activity.

*Vignette 2* (ACARA 2013c) is about a teacher reflecting on a lesson with students with significant intellectual and physical disabilities still acquiring basic communication skills. The lesson uses an age-appropriate Australian Curriculum Geography topic on landscape forms (Uluru) for Year 8 students (approximately 13–14 years old) to develop sensory and communication skills within the general capability Literacy Continuum. The lesson uses the physical resource of clay to create an Uluru shape and focuses on development of recognition and understanding of the phrase ‘who wants a turn’ in conjunction with the ‘turn’ sign, and the personal development of turn-taking. Some students have intentional communication. For others, she is working to build intention and association through repetition—‘I assign the meaning for now.’ Multiple goals are established through age-appropriate curriculum, and personal development and literacy capabilities. The extent to which these are shared with and understood by the students must again be inferred by the teacher. In this context of teaching and learning, the teacher says that knowledge of individual students is paramount—teachers ‘must understand the learner and who they are.’ The teacher also creates an implicit personal development goal through establishing positive bonds with the students. In the vignette, she looks around the students and asks ‘Who would like a turn?’ [with sign of hand turning]. When a student looks at her, she responds ‘Taylor you’re looking at me so I’m going to give you a turn’ [repeating hand sign with the word ‘turn’]. Further feedback to Taylor is provided by the teacher, while close to and looking into Taylor’s face, saying to the whole group: ‘I think Taylor likes the gritty feeling of Uluru.’ Taylor smiles in response. For another student, the teacher says ‘Good work Shane [thumbs up sign], I love the way you’re experiencing our Uluru.’ How do these interactions fit within AfL principles? Is the feedback being given task versus student oriented? Is it a combination? In this context do different principles of praise versus task feedback, as compared to those identified in ‘majoritarian’ principles (Ravet 2013, p. 954), apply to motivate ongoing learning? This is just one area where we need more research on effective AfL practices for all students.

## 4.5 Conclusion

Australian assessment policy, both through the Melbourne Declaration (MCEETYA 2008) and state and territory guidance, endorses teachers’ AfL practice for all students. Generic principles of AfL are referenced extensively. A question that arises is the degree of support provided to teachers in implementing both curriculum and these principles of AfL, particularly for diverse students, that is, in our discussion, students from diverse language and cultural backgrounds, and students with disability. The examples we discuss are clearly sited within Australian equity, curriculum, and assessment policy contexts at national, state, and territory levels. Our analyses of the teacher–student interactions in the four examples with students with disability show that the teacher practices are consistent with AfL

principles, even though AfL was not the primary focus of the examples. We also infer that these teachers' practices are informed and enabled by Australian curriculum and assessment policies and embedded school assessment practices. The Australian Curriculum age-appropriate alignment for equity in learning for diverse students that provided a new and challenging environment for the teachers and frames the learning goals in three of the four examples. We cannot, of course, generalise the practices of these teachers to all Australian teachers working with diverse learners, whether due to disability, culture, or language. However, we consider the broad Australian policy frameworks of equity, high expectations for all, and AfL should enable teachers to engage with AfL assessment principles for all students to promote learning.

The four examples involve a range of disabilities from a student with a learning disability affecting literacy, to students with various disabilities in inclusive mainstream education, to students who are nonverbal (cannot communicate by English language) and in some cases preintentional (cannot necessarily verbally or by sign indicate preference or intentions). The main finding that emerges from these examples using an AfL analysis framework is that while policy enables AfL practice, knowing the student is the primary informant of all elements of teachers' AfL practice. There is a gap between the broad expectations of policy and guidance, and the intensive teacher–student interactions of the classroom. Policy and official documents can never replace the expert teacher. However, the challenge for AfL is the creation of sufficient policy guidance and support to enable all teachers to move more rapidly on the pathway to expertise that integrates AfL practices with their knowledge of their students.

Our analyses of the examples in conjunction with the overview of research on AfL for diverse students highlight how much more research in this area is still needed. Our starting point is that if AfL is effective practice it should be enacted with all students. However, the examples presented in this chapter challenge the way generic AfL is currently advocated from the majoritarian perspective. We suggest the following guidance for policy, practice, and future research on AfL, especially for these students. A quandary in Australia continues to be who will have the responsibility for such research and development and support for the professional development of teachers. While common sense indicates that resources at the national level tied to the Australian Curriculum, as discussed in the two vignettes here, are most practicable, the need may be for policy, research, and practical guidelines to be developed at different systemic levels to enhance the work of teachers.

Firstly, policy makers should endorse, and researchers should undertake, further empirical research into effective implementation of AfL for diverse students, including diversity related to language and cultural backgrounds, and disability. In addition to examining the implications of currently advocated AfL principles for these students, such research should examine the implications of the implicit and explicit learning progressions that must inform AfL.

Explicit policy and resources are needed for teachers, and students and parents, to bridge the link between AfL and learning by students with diverse needs. These

should both draw on and enhance the empirical research evidence base for effective AfL principles and practices for these students. It is important that such policy and resources reflect the ‘spirit’ (Hayward 2014; Marshall and Drummond 2006) or ‘essence’ (Popham 2008) of AfL, given the need for teachers to consider each child as an individual. While some guidance is available, teachers could be given more support in the day-to-day planning of student learning through state-based initiatives such as the GIL.

Within schools, implementation of AfL for diverse students must go beyond generalised and generic principles to consideration in each specific context of the best approach to teaching and learning interaction, goal setting, and feedback for each child. Teachers need to consider whether the nature of assessment evidence and feedback are linguistically and culturally appropriate for each student. Teachers should implement AfL not only to scaffold student discipline learning but also to scaffold their induction into learning and classroom interaction discourses as appropriate.

Finally, but not least, following AfL principles, more investigation is needed on how diverse students can be successfully engaged in sensitive ways in responsibility for their own learning and the learning of peers, through self- and peer assessment. Technology such as tablets may be the critical new tools that enable these students to develop agency in their own learning and documentation of learning progress.

## References

- Abedi, J. (2010). Research and recommendations for formative assessment with English language learners. In H. Andrade & C. Cizek (Eds.), *Handbook of formative assessment* (pp. 181–197). New York, NY: Routledge.
- Albright, J., Knezevic, L., & Farrell, L. (2013). Everyday practices of teachers of English: A survey at the outset of national curriculum implementation. *Australian Journal of Language and Literacy*, 36(2), 111–120.
- Assessment Reform Group. (1999). Assessment for learning: Beyond the black box. Retrieved from [http://assessmentreformgroup.files.wordpress.com/2012/01/beyond\\_blackbox.pdf](http://assessmentreformgroup.files.wordpress.com/2012/01/beyond_blackbox.pdf). Accessed April 11, 2016.
- Assessment Reform Group. (2002). *Assessment is for learning: 10 principles. Research-based principles to guide classroom practice*. [http://assessmentreformgroup.files.wordpress.com/2012/01/10principles\\_english.pdf](http://assessmentreformgroup.files.wordpress.com/2012/01/10principles_english.pdf). Accessed July 20, 2015.
- Attorney-General. (2005). *Disability standards for education 2005*. Canberra, Australia: AG. <http://www.comlaw.gov.au/Details/F2005L00767>. Accessed July 20, 2015.
- Australian Bureau of Statistics. (2015). 4221.0—*Schools, Australia, 2014*. <http://www.abs.gov.au/ausstats/abs@.nsf/mf/4221.0>. Accessed July 20, 2015.
- Australian Curriculum, Assessment and Reporting Authority. (2012). *The shape of the Australian Curriculum* (Version 4.0). [http://www.acara.edu.au/verve/\\_resources/the\\_shape\\_of\\_the\\_australian\\_curriculum\\_v4.pdf](http://www.acara.edu.au/verve/_resources/the_shape_of_the_australian_curriculum_v4.pdf). Accessed July 20, 2015.
- Australian Curriculum, Assessment and Reporting Authority. (2013a). *An integrated approach* [Video]. <http://www.acara.edu.au/illustrations/IOPL00006/>. Accessed July 20, 2015.

- Australian Curriculum, Assessment and Reporting Authority. (2013b). *Student diversity*. <http://www.australiancurriculum.edu.au/StudentDiversity/Pdf/StudentDiversity>. Accessed July 20, 2015.
- Australian Curriculum, Assessment and Reporting Authority. (2013c). *Uluru: A landform in a landscape* [Video]. <http://www.acara.edu.au/illustrations/IOPL00001/>. Accessed July 20, 2015.
- Australian Curriculum, Assessment and Reporting Authority. (2015). *Personalised learning*. <http://www.australiancurriculum.edu.au/studentdiversity/eal-d-personalised-learning>. Accessed July 20, 2015.
- Baird, J., Hopfenbeck, T., Newton, P., Stobart, G., & Steen-Utheim, A. (2014). *State of the field review: Assessment and learning*. Report for the Norwegian Knowledge Centre for Education. <http://taloe.up.pt/wp-content/uploads/2013/11/FINALMASTER2July14Bairdeta2014AssessmentandLearning.pdf>. Accessed July 20, 2015.
- Bennett, R. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy and Practice*, 18, 5–25. doi:10.1080/0969594X.2010.513678
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, 5, 7–74. doi:10.1080/0969595980050102
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21, 5–31. doi:10.1007/s11092-008-9068-5
- Board of Studies (New South Wales). (2015). *Assessment for learning in the years 7–10 syllabuses*. <http://arc.boardofstudies.nsw.edu.au/go/sc/afll/>. Accessed July 20, 2015.
- Bourke, R., Mentis, M., & Todd, L. (2011). Visibly learning: Teachers' assessment practices for students with high and very high needs. *International Journal of Inclusive Education*, 15, 405–419. doi:10.1080/13603110903038488
- Carrington, S., & MacArthur, J. (Eds.). (2012). *Teaching in inclusive school communities*. Milton, Qld, Australia: Wiley.
- Colbert, P., & Cumming, J. (2014). Enabling all students to learn through assessment: A case study of equitable outcomes through the use of criteria and standards. In C. Wyatt-Smith & V. Klenowski (Eds.), *Designing assessment for quality learning* (pp. 207–227). Dordrecht: Springer.
- Crooks, T. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58, 438–481.
- Cumming, J. (2012). *Valuing students with impairments: International comparisons of practice in educational accountability*. Dordrecht: Springer.
- Cumming, J. J., & Mawdsley, R. (2012). The nationalisation of education in Australia and annexation of private schooling to public goals. *International Journal of Law and Education*, 17(2), 7–31.
- Department of Education and Communities. (2011). *Students with language backgrounds other than English in NSW public schools: 2011*. [https://www.det.nsw.edu.au/media/downloads/about-us/statistics-and-research/key-statistics-information/lbote\\_students.pdf](https://www.det.nsw.edu.au/media/downloads/about-us/statistics-and-research/key-statistics-information/lbote_students.pdf). Accessed July 20, 2015.
- Department of Education and Early Childhood Development. (2013). Assessment advice. Victoria: DEECD. <http://www.education.vic.gov.au/school/teachers/support/Pages/advice.aspx>. Accessed July 20, 2015.
- European Agency for Development in Special Needs Education. (2009). *Assessment for learning and pupils with special educational needs*. [https://www.european-agency.org/sites/default/files/assessment-for-learning-and-pupils-with-special-educational-needs\\_assessment\\_for\\_learning\\_en.pdf](https://www.european-agency.org/sites/default/files/assessment-for-learning-and-pupils-with-special-educational-needs_assessment_for_learning_en.pdf). Accessed July 20, 2015.
- Fuchs, L., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, 53, 199–208. doi:10.1177/001440298605300301
- Fuchs, L., Fuchs, D., Karns, K., Hamlett, C., Kataroff, M., & Dutka, S. (1997). Effects of task-focused goals on low-achieving students with and without learning disabilities. *American Educational Research Journal*, 34, 513–543. doi:10.3102/00028312034003513

- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81–112. doi:[10.3102/003465430298487](https://doi.org/10.3102/003465430298487)
- Hayward, L. (2014). Assessment for learning and the journey towards inclusion. In L. Florian (Ed.), *SAGE handbook of special education* (2nd ed., pp. 523–535). London, UK: SAGE.
- Hayward, L., & Spencer, E. (2014). Assessment for learning. In M. Carroll & M. McCulloch (Eds.), *Understanding teaching and learning in primary education* (pp. 162–175). London, UK: SAGE.
- Klenowski, V. (2014). Towards fairer assessment. *The Australian Educational Researcher*, 41, 445–470. doi:[10.1007/s13384-013-0132-x](https://doi.org/10.1007/s13384-013-0132-x)
- Marshall, B., & Drummond, M. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education*, 21, 133–149. doi:[10.1080/02671520600615638](https://doi.org/10.1080/02671520600615638)
- Ministerial Council on Education, Employment, Training and Youth Affairs. (2008). *Melbourne declaration on educational goals for young Australians*. [http://www.curriculum.edu.au/verve/\\_resources/National\\_Declaration\\_on\\_the\\_Educational\\_Goals\\_for\\_Young\\_Australians.pdf](http://www.curriculum.edu.au/verve/_resources/National_Declaration_on_the_Educational_Goals_for_Young_Australians.pdf). Accessed July 20, 2015.
- Perrenoud, P. (1998). From formative evaluation to a controlled regulation of learning processes. Towards a wider conceptual field. *Assessment in Education: Principles, Policy and Practice*, 5, 85–102. doi:[10.1080/0969595980050105](https://doi.org/10.1080/0969595980050105)
- Pitoniak, M., & Royer, J. (2001). Testing accommodations for examinees with disabilities: A review of psychometric, legal, and social issues. *Review of Educational Research*, 71, 53–104. doi:[10.3102/00346543071001053](https://doi.org/10.3102/00346543071001053)
- Popham, J. (2008). *Transformative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Powell, S., & Jordan, R. (2012). *Autism and learning: A guide to good practice* (2nd ed.). London, UK: Routledge.
- PricewaterhouseCoopers. (2013). *2012 trial of the nationally consistent collection of data on school students with disability* (Final report). Sydney, NSW: PricewaterhouseCoopers. [https://docs.education.gov.au/system/files/doc/other/attachment\\_13-2012\\_trial\\_of\\_the\\_nationally\\_consistent\\_collection\\_of\\_data\\_on\\_school\\_students\\_with\\_disability.pdf](https://docs.education.gov.au/system/files/doc/other/attachment_13-2012_trial_of_the_nationally_consistent_collection_of_data_on_school_students_with_disability.pdf). Accessed July 20, 2015.
- Queensland Curriculum and Assessment Authority. (2014). *Guide for individual learning*. Brisbane, QLD: QCAA. [https://www.qcaa.qld.edu.au/downloads/senior/qcia\\_guidel\\_ind\\_learn.pdf](https://www.qcaa.qld.edu.au/downloads/senior/qcia_guidel_ind_learn.pdf). Accessed July 20, 2015.
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28, 4–13. doi:[10.1002/bs.3830280103](https://doi.org/10.1002/bs.3830280103)
- Ravet, J. (2013). Delving deeper into the black box: Formative assessment, inclusion and learners on the autism spectrum. *International Journal of Inclusive Education*, 17, 948–964. doi:[10.1080/13603116.2012.719552](https://doi.org/10.1080/13603116.2012.719552)
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144. doi:[10.1007/BF00117714](https://doi.org/10.1007/BF00117714)
- Shaddock, A., Smyth King, B., & Giorcelli, L. (2007). *A project to improve the learning outcomes of students with disabilities in the early, middle and post compulsory years of schooling*. Canberra, ACT: Australian Government.
- Shute, V. (2008). Focus on formative feedback. *Review of Educational Research*, 78, 153–189. doi:[10.3102/0034654307313795](https://doi.org/10.3102/0034654307313795)
- Trumbull, E., & Lash, A. (2013). *Understanding formative assessment: Insights from learning theory and measurement theory*. San Francisco, CA: WestEd. [http://www.wested.org/online\\_pubs/resource1307.pdf](http://www.wested.org/online_pubs/resource1307.pdf). Accessed July 20, 2015.
- Van der Kleij, F. M., Feskens, R. C. W., & Eggen, T. J. H. M. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research*, 85, 475–511. doi:[10.3102/0034654314564881](https://doi.org/10.3102/0034654314564881)

- William, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching learning* (pp. 53–82). Mahwah, NJ: Lawrence Erlbaum Associates.
- Wyatt-Smith, C., Elkins, J., Colbert, P., Gunn, S., & Muspratt, A. (2007). *Changing the nature of support provision-students with learning difficulties: Interventions in literacy and numeracy project (InLaN)* (Vols. 1–4). Canberra, ACT: Department of Education, Science and Training, Australian Government.

# Chapter 5

## Formative Assessment Policy and Its Enactment in the Philippines

Patrick Griffin, Louie Cagasan, Esther Care, Alvin Vista and Fe Nava

**Abstract** In this chapter formative assessment is defined as a process of providing information to teachers to focus instruction on the improvement of student learning outcomes. The Department of Education in the Philippines in 2012 started the implementation of a new curriculum within a new structure which extended the education system from a Grade 1–10 to a K–12 structure to bring it into line with the developed world. A significant part of the curriculum reform is an assessment framework that includes formative approaches to assessment. The goal of the policy on formative assessment is to help teachers recognise relevant intervention practices that will improve student learning outcomes. An observation study of the link between assessment and teaching in a sample of 61 classroom lessons identified baseline practices and ways in which the emerging policy of the Department of Education in the Philippines could be promulgated. The observation study focused on classroom organisation, teacher instructional and assessment strategies, lesson structure, resources used by teachers, and student involvement in class work, as well as both formal and informal assessment practices. The observations were documented in a series of narratives aimed at identifying variation between teachers within grade level and disciplines. It emerged that a lesson structure which lingers from the previous curricular approach may be both the major inhibiting factor regarding formative use of assessment data and the most obvious opportunity for change.

---

P. Griffin (✉) · E. Care · A. Vista  
University of Melbourne, Parkville, Australia  
e-mail: p.griffin@unimelb.edu.au

L. Cagasan · F. Nava  
University of the Philippines Diliman, Quezon City, Philippines

A. Vista  
Australian Council for Educational Research, Camberwell, Australia

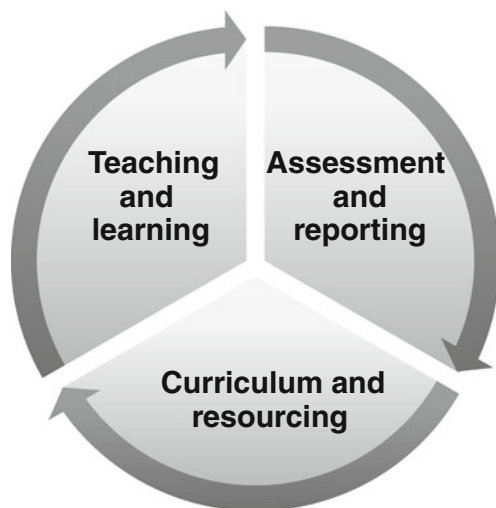


## 5.1 Introduction

The Department of Education of the Government of the Philippines implemented the Basic Education Sector Reform Agenda, known as BESRA, in 2012. In this country of over a hundred million people, any change is momentous. This particular reform is of major importance. It extends the education system from a ten-year to a thirteen-year system, in line with many nations globally and the rest of Asia, where the Philippines is the last country to make the move. It has reviewed and rewritten its curriculum. It has introduced mother tongue instruction in the first 3–4 years of education, revised its sequencing of teaching of mathematics and science in secondary education, and promoted an emphasis on applying understanding and skills rather than on content knowledge accumulation. The logistical, financial, and training implications of these reforms challenge the reformists and the country. The challenges reverberate to the level of the classroom, as will be seen in this analysis of formative assessment approaches in classes in Manila, the capital of the country.

The changes being implemented in the Philippines education system involve, at a minimum, three components. In any education program, there is a need to synchronise activities and philosophies across teaching and learning strategies, assessment and reporting procedures, and curriculum and resourcing (Fig. 5.1). If change is to be introduced at a classroom level, these three components need to be adjusted at a system level, and they need to inform and drive policy change. If all three components are not changed, the chance of introducing sustainable improvements in the classroom is diminished. In the Philippines, the first emphasis in the reform program has been on rewriting the curriculum. This raises immediate issues around its associated component of resourcing. It also raises issues of pedagogy and of assessment in order to ensure that the objectives of the curriculum can be realised.

**Fig. 5.1** Interdependence of components



Curriculum reform is just one part of the change that is required to remodel an education system. New curriculum documents can contain the best information and a major shift in thinking about learning and development, but the documentation alone cannot manage the shift in schools and classrooms. This is especially true when the shift is about educational philosophy. The goal of the revised education system is to produce ‘holistically developed Filipinos who have 21st century skills and are prepared for higher education, middle-level skills development, employment, and entrepreneurship.’ The shift is from a teaching and learning emphasis on process and content to an outcomes or skills emphasis. In the Philippines, assessment practices have centred on assignment of scores and letter grades. Unless the curriculum shift is accompanied by an assessment and reporting change that requires skill levels and outcomes to be reported, teachers will interpret the change in idiosyncratic and unintended ways. Reporting grades and scores simply emphasises the importance of grades and scores. By contrast, reporting skills, developmental levels, and social, intellectual, and performance quality growth emphasises that skills are what really matters. The curriculum is described as enriched, learner centred, decongested, seamless, responsive, and technology enhanced. These characteristics and aspirations can encourage teachers to focus on generalised developmental learning for their students. In order to do so, practices in the classroom must reflect the educational philosophy that has coined the terms. These practices include teaching and learning strategies, emphasis on the reporting of skills and development, and use of assessment information to support teaching and learning attuned to the developmental approach.

In the context of education reform in the Philippines, formative assessment is seen as the use by the teacher of assessment to inform teaching interventions. It is well aligned with Black and Wiliam’s 2009 statement, in that an assessment functions formatively to ‘the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make 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’ (p. 9). Black and Wiliam’s definition of formative assessment contains some important elements. The first is that it is about interpreting the assessment information or evidence to make decisions about the next steps in instruction. It is about informing teaching in order to improve student learning. It does not exclude the teacher encouraging the students to be involved in their own assessment but instead emphasises the responsibility of the teacher to use assessments in many different ways to improve student learning. In many instances the use of a developmental framework of learning can actively encourage students to understand their learning trajectory and become increasingly involved in monitoring their own learning development. However, for most situations the student and teacher need to collaborate in the monitoring process and ultimately the teacher’s professional judgment is required. It needs to emphasise *assessment for teaching* (Griffin and Care 2014). Our rationale for describing the formative approach as assessment for teaching is discussed in the following paragraph.

Some of the more recent literature on formative assessment focuses heavily on the use of feedback between teachers and students. Feedback is a part of formative assessment, but the real heart of formative assessment is a teacher's capacity to use the data to make decisions about intervention and the type of feedback that might be given to students to help them learn. Feedback is often regarded as any information given to the student about their current performance and the proposed performance level to which they might aspire. At best, it compares the current performance with a desired performance and as such would be regarded as the heart of a needs assessment in learning and teaching. But pointing out gaps in learning does not help the student improve learning and certainly does not help the teacher improve teaching. Black and Wiliam (1998) started with a definition of assessment itself. They regarded assessment as all activities undertaken by teachers that provide information to be used as feedback to modify teaching and learning activities. 'Such assessment becomes *formative assessment* when the evidence is actually used to adapt the teaching to meet student needs' (p. 2). This model is at odds with Kahl (2005) who also linked formative assessment directly to teaching. However, Kahl defined it in a deficit framework as 'a "midstream" tool used to identify specific student misconceptions and mistakes while the material is being taught' (p. 11). This deficit approach is not compatible with a belief in developmental learning.

Heritage (2010) is clear that formative assessment is not a tool, but a process, and is implemented in the context of natural progressions of learning within domains of interest: 'teachers need to have in mind a continuum of how learning develops in any particular knowledge domain so that they are able to locate students' current learning status and decide on pedagogical action to move students' learning forward' (Heritage 2008, p. 2). This requires teachers to not only have pedagogical skills but to understand deeply the nature of their discipline and how learning develops with it (Hutchinson et al. 2014). Such a developmental model of learning draws on the work of Glaser (1981), who defined criterion-referenced interpretation frameworks as those which allow monitoring of progress through stages of increasing competence. The model also allows the zone of proximal development (ZPD) (Vygotsky 1978) to be recognised as one of the stages of competence where the student needs assistance to learn. Rasch (1960) showed how this stage could be identified with the learning indicators that define the ZPD located where the student ability is equal to the difficulty of the assessment task, and hence the student has approximately a 50 % chance of success. For a review of this developmental model, see Griffin (2007). The model assumes that students can be located on a developmental continuum that describes stages of increasing competence (Glaser 1981). If these stages are described by a cohesive collection of skills that a student has a 50:50 chance of demonstrating (Rasch 1960), then the skills provide a description of the ZPD. This is important information for the teacher in that it requires assessment activities undertaken by teachers (Black and Wiliam 1998) and that the teacher interprets the information in a developmental manner, using either an explicitly defined continuum or the teacher's own intuitive scale in

order to make decisions about how best to help the student progress to the next stage of competence. This does not rule out information provided by the student that the teacher can use, but it does demand a process of collection, interpretation, and decision making by the teacher to help scaffold student development to the next stage of learning. Feedback to the student then takes on a more developmental approach. The assessment activities on their own cannot do this. Hence, we use the expression ‘assessment for teaching’ rather than ‘assessment for learning.’

We consider that assessment for teaching overlaps with assessment for learning and formative assessment in their basic formats. It involves a collaborative teaching cycle in which teachers collaborate in teams to make decisions about targeted teaching practices that focus on scaffolding learning for individual students. It is represented in a five-step iterative cycle which includes but goes beyond Sadler’s (1998) conditions for implementing formative assessment:

1. What is the student ready to learn? (ZPD)
2. What is the evidence for this?
3. What are the possibilities for learning and which approach is best?
4. What are the criteria of success?
5. What is the evidence of success?

Assessment for teaching is applicable within a growth model and dependent on an understanding of the developmental progression being used as a criterion-referenced framework for interpreting the evidence of student location on the progression. This ensures three things: that student growth or learning has direction and is not a random collection of goals; that there is an order to the growth or a sequence decided upon by the teacher in collaboration with a colleague and, where appropriate, with the student; and that there is an amount of learning agreed on in terms of level on the progression. This could be a set of skills acquired but is almost never described in terms of score increase, which we consider to be meaningless in an instructional and learning centred growth model. As more fully explained and exemplified in Griffin and Care (2014), assessment for teaching:

1. is situated within a growth model of teaching and learning;
2. involves a collaborative process between classroom process stakeholders (teacher, peers and students where appropriate);
3. demands that the assessment and learning depict direction, order and magnitude (three of four properties of fundamental measurement);
4. requires evidence based decisions regarding instruction, scaffolding, and outcomes;
5. is an iterative process involving the five questions listed above.

Compared to AfL, assessment for teaching emphasizes the central role played by at least two teachers who collaborate regarding the above questions: the resources needed (including but not exclusively other students), the scaffolding strategy, and evidence of attainment. The teacher is central but the student is the focus;

scaffolding is the primary strategy but student learning is the goal. These were the issues under consideration when exploring classroom assessment practices in the Philippines.

## 5.2 Method

The project, from which the results reported here derive, was established to conduct a national survey of teachers in their practices in assessment. The purpose of the larger study was to identify the variation in practices between the following elements of the system:

- teachers within schools
- schools within districts
- subjects within schools
- grade levels within and between schools
- teachers within subjects, and
- regions or districts within the system.

The rationale for observing these differences was to provide data that would inform the linking of successful practices to student learning outcomes. Several assumptions were made under this contextual framework:

1. Measures of student performance would be available;
2. Sufficient variation within and between each of these elements would be available and measurable;
3. Teacher practices could be linked directly to student performance; and
4. Formative assessment existed in the schools.

In identifying the degree to which formative assessment practices can produce the outcomes anticipated of them, such measurable and verifiable data sources are essential. In beginning to examine these sources of variation in assessment practices, sixty-one classroom visits and observations were conducted. Summaries of these classroom observations and interviews with teachers were documented in narrative form and the narratives then explored for patterns to use in the third and fourth assumptions above. Lewin's (1947) force field analysis approach was used to examine the forces facilitating formative assessment and the forces blocking such an approach. The forces were then examined to identify relevant actions and recommendations.

Systematic classroom observations were conducted in the last quarter of the school year 2013–2014. The observations ranged from Kindergarten to Grade 9 (or 3rd year of secondary school where students are normally within the age range 13–14 years old). The observations had a focus on mathematics and English subjects. Table 5.1 presents the sample descriptions and the details of the schools. The

**Table 5.1** Descriptive statistics of participating schools by grade levels of classes observed

Level	Frequency			DepEd schools			UP integrated school		
	Total	Math	English	No. of classes	Average class size	Average duration (min)	No. of classes	Average class size	Average duration (min)
Kind.	1	1	0	1	21	60	–	–	–
1	10	5	5	8	29	49	2	25	72
2	6	3	3	2	42	60	4	–	82
3	8	4	4	6	38	67	2	34	75
4	11	6	5	9	41	56	2	31	75
5	11	5	6	9	41	79	2	33	71
6	8	4	4	6	47	51	2	29	73
7	2	1	1	2	49	65	–	–	–
8	2	1	1	2	44	64	–	–	–
9	2	1	1	2	53	56	–	–	–
Total	61	31	30	47			14		

Note Mean values are rounded to the nearest decimal place

year levels, subject observed (English and or mathematics), and typical class sizes and lesson duration across government and independent schools are presented. Twelve public schools in Quezon City, in Manila, were visited. Most schools were under the supervision of the Department of Education (DepEd) ( $n = 11$ ) and one was a laboratory school under the supervision of University of the Philippines (UP). In terms of class size, Kinder and Grade 1 levels for Department of Education schools were smaller in number than other grade levels. For the UP school, the class sizes were within the range of 25–35 students. The duration of a class session in a DepEd school is typically sixty minutes, and for UP the suggested length of class period is 75 min. The average class duration observed indicates that most classes are near the mandated length.

Table 5.2 displays the summary of class statistics of the Department of Education schools observed. Column 3 shows that the class sizes increase with the

**Table 5.2** Descriptive statistics for Department of Education classes observed

Department of Education levels	Number of classes	Average class size	Average class duration (min)
Basic Education: Kinder	1	21	60
Basic Education: G1–G6	40	39	62
Secondary Education: G7–G9	6	49	62
Total	47	40	

Note Mean values are rounded to the nearest whole number

grade level of students in these particular schools, all of which were in Quezon City, which covers about a quarter of metro Manila, and is the most populated city in the Philippines.

The classes generally start with checking of work done, identification of correct responses and of those students who achieved these. When the teacher asks for responses that are a matter of direct recall, she receives more responses than when she moves to definitional questions. The teacher then models an activity with some student participation, and then sets a task for the students to complete as groups. Through to the completion of this activity, not all students participate or are included in the work. In the next activity all students are included, and the teacher identifies winning groups as those who complete the activity first. All groups are given the opportunity to present their results. Requests to the students to generalise their understanding are met with few correct responses. The teacher concludes by distributing a 5-item quiz to students, which the students then peer mark. The teacher calls out the correct responses, and then asks for students with a pass grade to self-identify. Those students lower than pass are instructed to study the activity.

## 5.3 Results

### 5.3.1 *Classroom Practices Through a Formative Assessment Lens*

**Common structure:** The narratives of the 61 classroom observations provide a persistent impression that there is a common structure for the lesson plan. The formats used for lesson delivery appeared to be inflexible. The lesson began with a class activity, followed by small group activity, question-and-answers with show of hands, and then a summary session. Sometimes the class activity consisted of a recitation activity at the beginning of the class and at other times the content recitation was performed by individual students at the front of the class but only if they were prepared to demonstrate that there was a chance they did not understand what was taught. What may have been an attempt at formative assessment may have been counterproductive and provided signals that there is a need for teacher development in the use of data to promote learning for individual students. The primary focus of the teacher is on identification of correct responses and correct responders.

**Student Behaviour:** Results from assessments were sometimes used to modify teaching strategies for an entire class. There was little evidence that assessment results were targeted to individual students or to small groups of students. Individual intervention was often based on behaviour (e.g., low interest, truancy), rather than on level of skill as indicated by assessment. Teachers used student behaviour to cue them on whether students were paying attention, were uninterested or not motivated, were understanding the lesson or becoming confused, and were

mastering the lesson. Many teachers judged the extent to which students knew or had learned the lesson content by the way the students reacted rather than by a formal assessment. For example, mastery was assumed of students who demonstrated behaviours associated with understanding—raising hand during question and answer; apparent attentive listening; participation in discussion; asking questions. The use of such indicators and anecdotal evidence can be valid, but few teachers understood that both responses and nonresponses could be used formatively. Few marked or noted students' outputs, or if they did, they did not record results as part of the final grade. Many teachers, however, did incorporate results of the routine class quiz as part of the evaluation of student performance.

Generally, when the teacher asked questions, a show of raised hands was used by the teacher as an index of understanding at the class level. However, the teachers did not demonstrate realisation of the potential of the assessment information and did not maximise the effectiveness of the strategy. In part, this may be attributed to the preponderance of low order questions, requiring mainly closed or recall responses.

**Assessment:** Assessment appeared to be uniform for every student in the class, and results were interpreted in aggregate form, such as what percentage of the class was above or below a given threshold. Occasionally, the teacher interacted with individual students but this was limited and did not appear to be connected with the use of data to inform intervention and direction of student learning. A quiz was typically held at the end of the class. This consisted of approximately five to eight true/false or short-answer, supply-type items. Students scored their own test answers or exchanged tests in order to have a partner score them. The teacher did not get involved in scoring individual student results on quiz questions. The teacher occasionally collected the student books or test sheets and recorded the performance on these quizzes at an individual student level. However, there was no apparent use made of the data.

**Grouping:** Grouping students in the class was typically used to encourage learning through competitive participation and engagement. Some teachers appeared to be successful in using this approach for engagement purposes. Students enjoyed, or appeared to enjoy, the competitive nature of some of the group tasks. The actual grouping appeared to be arbitrary or based on subjective judgment on ability.

**Curriculum pressures:** Many teachers were aware of, and sensitive to, the amount of time taken by various activities in the class. The perception that the curriculum is difficult to cover in terms of the breadth of scope and sequence may account for the formulaic approach to teaching that is adopted. The formulaic nature of the lesson appears to achieve compliance with curriculum objectives on the part of the teacher but does not necessarily achieve learning on the part of the student. The classroom is an environment in which the teacher must deliver the set curriculum topic, rather than a venue where the focus is on the student learning. Content varied from teacher to teacher and was subject to school level decisions.



**Students:** There was little evidence of students seeking formative feedback from the teacher and very little evidence of the kinds of strategies that the students use in their efforts to learn. Students did not expect formative feedback regarding their learning outcomes. The quiz at the end of the class was commonly the only method of providing learning information or feedback to students through identification of correct/incorrect responses. During classes teachers sometimes asked if there were any questions, but in some classes this practice tended to discourage the students, who appeared to be intimidated due to the consequence of being called to the front to have their question explained in front of their peers.

**Summary:** While teachers taught different content within subjects, and for the most part schools practised a fair amount of autonomy in terms of the delivery of content, there was little or no autonomy in lesson design, structure, and format, or assessment practices. It is clear that the lesson structure (see Table 5.3) is associated with:

1. Summative assessment at the end of the lesson
2. Pressure on teachers to ensure that the content of the curriculum is covered in each lesson
3. Limited individual feedback or feed forward for students
4. Pressure on teachers to monitor their own performance but not that of the students
5. Common format.

Insofar as assessment was practised in the classrooms, minimal emphasis was given to formative assessment and most assessment was summative and recorded as scores or percentage correct. The assessment skills of teachers resulted in a state of equilibrium in the classroom brought about by the pressure of an intense content-focused curriculum, large class sizes, a formulaic lesson plan and structure, regular and mistargeted assessments, and a lack of accountability for student learning. These observations made it clear that there was little chance of pursuing the original intention to survey the influence of variability between teaching strategies across grade levels, across subjects, across education levels, and across schools. Given the invariance of lesson structure and pedagogy, improvements in student learning were more likely to be a result of other factors. However, the observation study did provide evidence of a need for changes in pedagogy if formative assessment is to be used to improve teaching and learning.

In summary, the classrooms were environments in which:

1. Teachers asked questions and used a show of hands as an indicator of learning success.
2. Teachers sometimes collected student's notebooks, but provision of written comments as feedback was unusual.
3. Teachers sometimes set group projects and group work, as well as individual assignments.

**Table 5.3** Blocking forces

Blockers	Effect	Action to reduce the effect
B1 Formulaic lesson plan structure (teaching guides are given to the teachers) [Teacher interviews]	The focus on following a predefined structure results in less attention in determining how learning is developed in the classroom. In the implementation of the K–12 curriculum, the teachers are now advised to use daily lesson logs; however, many teachers prefer to stay with previous ways. The emphasis is on covering the curriculum as a priority	Shift the attention of daily lessons from activities and predefined structure to assessment of learning and assessment for teaching; lesson plans should include questions drawing out the inputs generated by the teacher. Lesson plans should include specific indicators that would inform the teacher about the current level of each student in order to provide assistance. Advice is needed for using quiz data at the beginning of each class to assist teaching
B2 Lack of teacher training in formative assessment [Teacher interviews]	The types of questions asked do not link to a skill progression. It is difficult for teachers to determine the development of the student in relation to a particular skill. Also, collected quiz data are not used to improve student teaching. Teachers are unclear about the purpose of assessment, and as a result, the link between assessment and teaching is not clearly established	Introduce to teachers the concept of skill progression; help teachers to emphasise the importance of skill development rather than content-driven perspectives; reform current perceptions that more activities mean more learning; help the teachers to use data from different types of questions to identify increasing skill levels
B3 Pressure of curriculum coverage [Teacher interviews]	Even if teachers know that there is a percentage of students who do not understand the lesson, they feel an urgency to proceed to the next lesson because they are expected to cover a set range of topics. This leads to some students not developing the skills or foundation knowledge needed to understand the next lessons	The curriculum should shift towards a developmental progression framework to reduce the pressure of curriculum coverage. The best option would be to focus on the foundation skills (cutting across different content areas) needed for students to go to a higher level. Emphasise skills in the curriculum rather than content
B4 Emphasis on summative assessment [Observation and narrative records]	Collection of data is used for the purpose of generating ‘grades’ which are used to represent the overall performance of the student. Attention given to the actual skills acquired by the students is minimal. With focus on summative assessment, areas that need to be improved are not identified	Assessment can focus on the developmental skills as a shift from content-based to a skill-based. Materials on skills and scaffolding need to be provided for teachers, perhaps as part of online resources to maximise reach and opportunity to access

(continued)

**Table 5.3** (continued)

Blockers	Effect	Action to reduce the effect
B5 Lack of a data management system at class and school levels [Observation and narrative records]	Not all the activities and outputs of the students are recorded consistently and systematically. As a result, these data are not used to develop plans for improving teacher practices and student learning	Within-class differentiated instruction can be implemented if targeted materials are available and teachers have reliable records of student progress. The resourcing issue can be addressed through provision of basic recording devices to teachers and different worksheets for groups of students. DepEd's current initiatives for an integrated information system may assist this
B6 Systemic focus on grades [Teacher interviews]	Related to 'Emphasis on summative assessment,' the documentation inside the classroom produces grading. However, the focus on student's level of understanding and skill is not reflected. If the student has a passing grade, he or she is eligible to move to another grade regardless of actual understanding	Instead of using an overall index of student performance, encourage teachers to employ different indices of student skill as part of recording and reporting
B7 No clear link between assessment results and instructional planning [Observation and narrative records]	Empirical data are not used to support teaching strategies to improve student outcomes	Provide teachers with materials and methods that assist them in collecting, analysing, and reporting student achievements and skill acquisition
B8 Heavy teacher workload [Teacher interviews]	Limits additional interventions that can be asked of teachers	Focus on improving materials on basic skills needed for teachers and on reducing teacher workload by using technology to ease labour-intensive routine tasks. Online resources may be needed
B9 Need for more various forms of accountability in the system and clarification of link between performance and compensation [Teacher interviews]	Accountability for test scores and a system of compensation linked to test scores encourages teaching to the test and concentration on improving performance in external tests	Provide assessment materials that are evidence based, in that they are linked to student skill development. Develop a compensation system that rewards teacher collaborative teams for collective student learning and skill development. Online training modules may be necessary

4. Records of student learning gains consisting of scores or grades are not well suited to planning instruction. Recorded descriptive information about progress of individual students in mark books was unusual. Anecdotal records were not observed at all. Most of the recordings appeared to rely on memory and in large classes these were problematic. Few records provided information about the class, the student, the topic, the syllabus, progress of students, or progress of the class, notwithstanding rigorous completion of marking spreadsheets.
5. Teachers recorded their own progress in terms of content coverage and lessons completed. Very few made evaluative or assessment comments about their own performance.
6. The procedures for teachers to synthesise information at the end of term in order to provide a report are clear and are based purely on summative data.
7. Records to students for purposes of instructional feedback, as opposed to feedback of grades, were non-existent.
8. Possibilities for collaboration with students or with colleagues were non-existent given the mandatory rate of coverage of the curriculum.

Accordingly, it was decided that the narratives would be analysed using Lewin's force field analysis to identify the operating forces within classrooms that might encourage or discourage the use of formative assessment.

### ***5.3.2 Force Field Analysis***

The force field analysis provides a base for Lewin's (1947) three-stage theory of change: unfreezing the existing equilibrium (disruption), moving towards the desired change (change forces), and then freezing the change at the new level (institutionalization). In this case it is necessary to find a way in which formative assessment could be part of the new equilibrium and put a system in place to support this equilibrium. The introduction of formative assessment will depend on changes to the resisting or negative forces and an enhancement of the facilitating forces. From the records of school visits, the elements of the force field analysis emerge. Table 5.3 presents the force field analysis for the blockers while Table 5.4 presents the analysis for the facilitating forces. Both tables list the forces, the effect of each on formative assessment practices, and the recommended actions either to weaken the blocking or strengthen the facilitating forces.

It should be noted that the suggested actions are drawn from local understandings of what might be possible to implement, rather than recommending actions that are too far beyond the current capacity of the system and its teachers.

**Table 5.4** Facilitating forces

Facilitating force	Effect	Action to enhance effect
F1 Student response to feedback [Observation and narrative records]	Feedback to students will encourage teachers to focus on what was taught and what is learned	Provide materials that help with systematic data collection and interpretation at the student and class level. Provide user-friendly materials that link assessment data to teaching strategies and student outcomes
F2 Policy shift towards formative assessment [DepEd Order No. 73, Series 2012 (DepEd 2012)]	The Department of Education is actively shifting focus to encourage the use of formative assessment in the classroom	Publicise the support from DepEd for a review of current assessment practices and encourage, disseminate, and reward implementation of formative assessment practices
F3 The need to improve student skill development [National Achievement Tests and TIMSS 2008 results (Arora et al. 2009)]	There is a need to improve awareness of international benchmarks. This has the effect of increasing the motivation within the education system to adopt new ways to improve student outcomes	Provide clear evidence that formative assessment has significant impact in improving student outcomes, both by improving teaching and by linking assessment with planning
F4 Potential of new technologies to aid teachers and ease their administrative workload, while also assisting in more systematic record keeping and data analysis [DepEd ICT4E Strategic Plan (DepEd 2011)]	Using technology in the classroom reduces teacher workload and increases the efficiency of record keeping. Technology also enables more accurate data analyses which then results in faster feedback to students and stakeholders	Provide online tools and proper training linking assessment and teaching, and assuring ease of analysis and interpretation of assessment data
F5 Teachers follow system-wide instructions closely [Teacher interviews]	Systemic interventions can be implemented uniformly across the system	Disseminate materials for teachers to help improve their formative assessment and reduce workload

*Note* Text in brackets indicates data source of the forces identified

## 5.4 Recommendations

From the observations of classes, it was clear that the majority of classes are run to a standard structure, with emphasis on moving through curriculum content. The focus is clearly on what the teachers do. The context for this is a highly centralised and hierarchical system in which authority relationships are preminent. The K–12 reform provides a major opportunity for changes in pedagogy and assessment.

Impediments to these changes lie primarily in the prescribed nature of how teachers and schools function, in lack of materials resourcing, in crowded classrooms, and in relatively low levels of teaching expertise.

The Department of Education was moving on these fronts, and two years into the reform process, was focussed on the national assessment framework—at system, school, and classroom levels. Early initiatives include train the trainer activities to develop the expertise of teachers in their assessments of literacy and numeracy in the early years. Similar initiatives in formative assessment are justified. At the most basic level, focus on questioning techniques to stimulate student thought and analysis is a reasonable first step to enhancing teacher skills. In the observations, questions measuring higher order thinking skills or critical thinking were very rare, and occurred, if ever, only in the form of reasoning, explaining, or defending an answer. Teachers need to be supported in framing questions both for assessment and for teaching purposes. There is a need for open-ended questions, diagnostic questions, information-seeking questions, challenge questions, action questions, questions on priorities, prediction questions, hypothetical questions, questions of extension, and generalisation questions in all subjects and at all grade levels. This change alone would have an important impact on the use of formative assessment and would blend assessment with teaching. The current practices in classrooms are reflective of a belief in the primacy of a correct response and of a competitive environment. Developing the teachers' skills in feedback strategies linked to their improved questioning strategies would be an important step in improving teachers' use of formative assessment strategies.

It is pragmatic to work within existing paradigms if these can offer the opportunity to reach the desired outcomes. The training which will be required of teachers to implement the reform will not take place as rapidly as the initial reform of curriculum process. It makes sense to identify teaching strategies which can be integrated within current practice, but there is also a need to focus on those practices which will reframe teachers' understanding of the learning process. Zuzovsky (2013) has drawn attention to the differential impact of particular teaching strategies across countries (or classrooms) which vary in educational performance. In relating classroom practices to educational outcomes in mathematics and science on large scale assessments, she found that some traditional practices, such as the use of a short quiz, had negative associations with achievement in lower performing countries, while some constructivist modes of instruction, such as students explaining their answers, had strong positive associations. This finding alone can inform how a 'mandated' component of the classroom could perhaps be turned to positive effect with some rethinking of how the quiz can be managed to act as formative assessment.

Arising from the force field analysis, three main educational components are highlighted for promotion of positive changes in educational assessment and consequently in quality of education. These are the availability of materials, the assessment system, and a technology platform for delivery of assessment and materials. These strategic components can provide the infrastructure to support teachers in their professional development in assessment and pedagogy (Table 5.5).

**Table 5.5** Recommendations, resources needed and timeline

Recommended action	Resources needed	Timeline
Change in formulaic lesson plan	Support from DepEd in developing and embedding successful formative assessment techniques in lesson plans, model lessons and exemplar assessments, questioning techniques and quizzes	Short term, 2 years
Training to provide better task-oriented feedback	Announced support from DepEd to assist agencies to develop online professional development sessions	Short term, 2 years
Training on test construction and development	Announced support from DepEd to assist agencies to develop online professional development sessions	Short term, 2 years
Shifting the focus of assessment—focussing on skill rather than content	Provision of exemplar materials and advice on how to use assessment data to help teaching	Short term, 2 years
Improvement in reporting to school administration	Provision of technologies to facilitate assessment interpretation, and implications for teaching and reporting	Short term, 2 years
Accountability for teaching and student results	Provision of enabling technologies to facilitate systematic data collection and reporting and collaboration among teachers	Medium term, 5 years
Improvement in reporting to parents	Provision of technologies to facilitate electronic reporting to enable reports of skills developed as well as grades (as a transition strategy)	Long term, 10 years

## 5.5 Conclusion and Implications

In a study of Philippines classroom assessment it became clear that the ability of most teachers to synthesise information about the students' levels of development, as well as their discrete skills and content learning, was not manifested in the delivery of lessons. Assistance for teachers to adopt an understanding of developmental approaches to learning and teaching is critical for the successful implementation of the K–12 education reform. Understanding of the interdependence of teaching and learning, of assessment and reporting, and of curriculum and resourcing is critical at this stage in the Philippines' reform process and is clearly visible through this analysis of the issues faced by teachers and students in the classroom. Pedagogical and assessment skills must be enhanced to support an educational philosophy of learning for all and achievement for all.

There are also several other lessons to be drawn from these observations. Firstly, the study highlights the need for formative assessment *of* and *by* teachers. Formative assessment of teachers can be addressed in many systems by reflecting on the development of basic pedagogical strategies. Questioning, for instance, will

become an important pedagogical skill in the 21st century. As the knowledge society develops, even in developing economies, education will be under increasing pressure to change. Teachers will be less able to be transmitters of information, and the role of a teacher will need to change. Questioning has always been an important skill, but its importance will increase as the skill of addressing questions to students replaces the didactic approach of telling. However, there may be cultural influences at work here; an effective approach to changing teacher questioning will be difficult. In a study of Vietnamese Primary teaching practices (Griffin et al. 2006), it was clear that teacher expertise was an important issue and questioning was used predominantly as a means of checking whether the students had attended to the teachers' knowledge transmission.

An important lesson from this research that could be of interest to an international audience is that the assistance to teachers is based on a detailed study of how they actually teach and assess students, and the recommendations made to the Department of Education were based on identification of barriers and facilitating factors. In short, the department was encouraged to implement a policy which takes teachers from where they are instead of where they should be. This issue may arise in many cultural settings where teachers are respected and rewarded for their knowledge. Paradoxically, changing these school cultures to enable more emergence of a knowledge society will be exacerbated by the very respect that is afforded to teachers as content experts. Griffin et al. (2006) and others reported that it was only recently (1999) that teachers were permitted, after an inspector reviewed their practices, to depart from the mandated Vietnam lesson plan. If this were to happen in the Philippines the country might founder in its stated goal of developing holistically prepared adults with twenty-first century skills. Ironically, the development of the twenty-first century Filipino requires pedagogical skills among teachers, starting with questioning strategies, flexible lesson plans, and formative assessment practices focused on their own learning as well as that of their students. The change in assessment practices needs to focus first on how teachers use assessment data to adjust their pedagogy to help student learning, and this of course includes using student involvement in the process.

## References

- Arora, A., Foy, P., Martin, M. O., & Mullis, I. V. S. (Eds.) (2009). *TIMSS and PIRLS advanced 2008 technical report*. International Association for the Evaluation of Educational Achievement (IEA). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Black, P. J., & Wiliam, D. (1998). Inside the Black Box: Raising standards through classroom assessment. *Phi Delta Kappa*, October 1–13.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31.



- DepEd. (2011). *Five-year information and communication technology for Education Strategic Plan* (DepEd ICT4E Strategic Plan). Manila: Department of Education, Republic of the Philippines.
- DepEd. (2012). *Guidelines on the assessment and rating of learning outcomes under the K to 12 basic education curriculum* (DepEd Order No 73, s. 2012). Manila: Department of Education, Republic of the Philippines.
- Glaser, R. (1981). The future of testing: A research agenda for cognitive psychology and psychometrics. *American Psychologist*, 36, 923–936.
- Griffin, P. (2007). The comfort of competence and the uncertainty of assessment. *Studies in Educational Evaluation*, 33, 87–99.
- Griffin, P., & Care, E. (2014). Assessment is for teaching. In P. Griffin (Ed.), *Assessment for teaching*. Melbourne: Cambridge University Press.
- Griffin, P., Nguyen, T. K. C., Gillis, S., & Mai, T. T. (2006). An empirical analysis of primary teacher standards in Vietnam. *Planning and Changing*, 37(1& 2), 71–92.
- Heritage, M. (2008). *Learning progressions: Supporting instruction and formative assessment*. Washington DC: Council of Chief State School Officers.
- Heritage, M. (2010). *Formative assessment and next generation assessment systems: Are we losing an opportunity?*. Washington DC: Council of Chief State School Officers.
- Hutchinson, D., Francis, M., & Griffin, P. (2014). Developmental teaching and assessment. In P. Griffin (Ed.), *Assessment for teaching*. Melbourne: Cambridge University Press.
- Kahl, S. (2005). Where in the world are formative tests? Right under your nose! *Education Week*, 25(4), 11.
- Lewin, K. (1947). *Resolving social conflicts: Selected papers on group dynamics*. New York: Harper & Brothers.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Neilson & Lydiche.
- Sadler, D. R. (1998). Formative assessment: revisiting the territory. *Assessment in Education*, 5(1), 77–84.
- Vygotsky, L. S. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.
- Zuzovsky, R. (2013). What works where? The relationship between instructional variables and schools' mean scores in mathematics and science in low-, medium-, and high-achieving countries. *Large-scale Assessments in Education*, 1(2), 1–19.

# Chapter 6

## Communication and Collaboration: The Heart of Coherent Policy and Practice in New Zealand Assessment

Jenny Poskitt

**Abstract** Educational policy formation is the result of multiple influences, locally and internationally, with power struggles amongst political, economic, academic, practitioner, and local cultural contexts. Various theories abound regarding the international flow and influence of educational policies, such as ‘borrowing,’ ‘lending,’ ‘travelling,’ and ‘knowledge transfer.’ This chapter argues that ‘policy adaptation’ is the more pertinent term because policies generally mutate from one setting to another. Two examples of New Zealand policy enactment in formative assessment, one related to National Curriculum Exemplars and the other to National Standards, illustrate the centrality of clear communication and collaboration across all players in the educational sector, as well as responsiveness to the local political, economic, and cultural context for coherent policy and practice to occur. The examples demonstrate the role all players can and need to take in shaping policy. Indeed coherent policies that meet the greatest challenge of all—implementation in practice—can only occur through extensive collaboration and communication.

### 6.1 Introduction

Learners are the ultimate benefactors of well-written and enacted educational policy, but the gap between policy writers’ intentions and practice in the classroom can be wide. Policy formation emerges in arenas at considerable distance from the classroom, whereby practitioners may be expected to implement a policy without awareness of the rationale or decision-making process. Part of the mystery is due to educational policy formation being the subject of multiple influences and power

---

J. Poskitt (✉)  
Massey University, Palmerston North, New Zealand  
e-mail: j.m.poskitt@massey.ac.nz

struggles amongst political, economic, cultural, and educational spaces (Clarke 2014; Waldow 2012). Endeavours to understand and clarify the process of policy formation have resulted in several theories such as ‘borrowing,’ ‘lending,’ and ‘travelling policies,’ but ‘policy adaptation’ appears to be a more appropriate term because policies generally mutate from one setting to another.

This chapter examines processes used to develop, disseminate, and implement assessment for learning policy in New Zealand through two examples: National Curriculum Exemplars, where multiple factors aligned and policy development was harmonious, and National Standards, where considerable contestation and struggle occurred. Communication and widespread collaboration amongst key players in the educational sector were pivotal factors to ultimate implementation at school and classroom level. The discussion turns firstly to theories of policy formation, then to principles of effective collaboration, and finally to their manifestation in two New Zealand examples of assessment policy enactment.

Concerning the terminology used in New Zealand, ‘formative assessment’ was the original term and typically appeared in more formal settings, such as early national policy documents and academic writing. In order to be more appealing to practitioners and to encourage them to embed assessment in their practice, the term ‘assessment for better learning’ was adopted in some documents (e.g., Department of Education 1989). Around 2000 ‘assessment for learning’ (AfL) became more popular (e.g., Absolum 2010), aligned with international trends.

## 6.2 Policy

Underpinning the notion of policy is an awareness of interconnected relationships and complex webs of influence that cause ebbs and flows of resistance, complicity, and advancement in processes, systems, beliefs, and practice. Generally, influence flows from those endowed with greater power, knowledge, financial, and people resources and who have garnered legitimation for their cause, to those of lesser endowment (Bridges 2014; Froese-Germain 2010; Steiner-Khamsi and Waldow 2012). This is the case whether policy is international, country, region, or school based (Adie 2014). Nevertheless, the influence is not a tidy linear process. Instead, policy formation is generally a messy process that occurs in a space of intense negotiation, reconciliation of competing views, and legitimation (Waldow 2012) that may become a force in itself. ‘When faced with problems requiring solutions, proposed policies must confront the tribunal of competing social forces from rival sectors and interest groups’ (Clarke 2014, p. 30). Although politicians formulate policy ideas for government, it is the policymakers who initiate the policy journey with administrative decisions and provide rationale to persuade practitioners of its value, so they have sufficient interest and motivation to implement it. But with educational policies, parents and potential employers also have a stake, being concerned about ultimate preparedness for the workforce (Clarke 2014).

### 6.2.1 Policy Influences

Education does not occur in a vacuum, for it is embedded within its country's cultural, economic, and political context; all of which influence and constrain the other. In an era of neoliberalism, economics have had a significant influence on education. Economic language has become commonplace within official education documents [e.g., inputs, outputs, outcomes, evidence-based, performance indicators, measurable, accountable, competitive, world class (Froese-Germain 2010; Steiner-Khamsi 2012)], and 'devoid of terms such as "caring relationships," "moral and ethical purposes of education," and "critical thinking"' (Froese-Germain 2010, p. 5). Within this socio-political milieu, education has been reframed as central to strengthening economic competitiveness through the greater influence of policy levers over finances in instigating change (Froese-Germain 2010). Moreover, in an age of increasing technological advancement, there has been a growing awareness of, and connection with, nations around the world. With easier access to information have come increased competition and the trend of modification of other nations' policies. Indeed, 'some authors fear we are abandoning our idiosyncratic conceptions of "good education" and are gradually converging toward an "international model of education" (Steiner-Khamsi 2012, p. 1).

Taking on board policies of other countries is often referred to as 'borrowing' and 'lending' policies (further examples of the infiltration of economic terms in education), as well as 'travelling' and 'embedded' policies. The 'lender' of policy is generally the nation(s) considered to have superior or more effective policies than the 'borrower' nation(s) receiving the policy (Steiner-Khamsi 2012). Such practices were common in colonial times and in the establishment of schooling, but unfortunate connotations remain of offering technical and professional support associated with challenging or replacing cultural norms, values, or ways of living, especially from white western origins to indigenous cultures (Bridges 2014). Implicit can be notions of power, imposition of cultural values as well as political and economic blackmail when policy borrowing is accompanied by financial aid (Bridges 2014; Froese-Germain 2010). Moreover, subtle and more overt pressures to 'raise educational standards' and to adopt 'best practices' or 'international standards' establish conditions of obligation to modify the educational system to be more aligned to those of other countries.

Some authors use more apparently neutral terms such as 'knowledge transfer' (Ozga and Jones 2006). Divala (2014) argues the more appropriate terms are 'policy transfer,' 'policy travelling,' and 'policy learning,' which have lesser connotations of power and obligation inherent in 'borrowing' and 'lending.' Travelling policy is defined as 'policy shaped by globalizing trends in pursuit of successful competition in the new knowledge economy,' while 'local inflections of policy are understood as embedded policy, mediated by local contextual factors that may translate policy to reflect local priorities and meanings' (Ozga and Jones 2006, p. 1). The current chapter maintains none of these terms express the iterative process of policy development, and argues instead for 'policy adaptation' since policies are modified,

indeed mutated, wherever they are implemented in accordance with local circumstances, economic, political, and human capacity. Such adaptation is discerned in the New Zealand examples below where the resulting policy is generated from negotiations and power struggles that occur amongst the policy players and enactors.

The thesis is somewhat supported by Ozga and Jones (2006) who argue the importance of ‘the local in responding to, and mediating globalizing pressures and travelling policies’ (p. 14). They further posit, ‘policies may get re-contextualized and remodelled according to local and natural histories, traditions and social relations, even when they are concerned directly with serving the knowledge economy, and even when they are apparently designed to ensure commodification and alienation of knowledge’ (Ozga and Jones 2006, p. 14). Further advancing these notions, Adie (2014) argues that the dialectical process is influenced by available resources and by the progression of participants from imagined to real change—through interaction with other participants and through negotiation of understandings and practices. Implicit is the understanding of co-constructing and developing ownership of a policy before implementation can occur by practitioners. While the pervasiveness of globalisation needs to be acknowledged, opportunities arise to modify global policies to select pertinent elements in accordance with the uniqueness of the particular local setting. Key underlying factors are the degree and nature of communication and collaboration amongst various players.

### 6.3 Collaboration

Successful collaboration requires: establishing awareness of and respect for the particular expertise of participating partners, inclusion of a diverse mix of strengths, a shared language and mutual learning, a growing sense of trust and friendship, as well as common goals and commitment towards achieving them (Elliott and Woloshyn 2001; Schutz et al. 2001). Working collaboratively can benefit members by creating a collective sense of belonging and being valued, by professional stimulation through interactive dialogue and exchange of ideas, and by the synergistic effect of accomplishing more together than can be accomplished alone. In collaborative situations, members may take risks within the group and break new ground that may not be possible alone (Elliott and Woloshyn 2001). However, collaborative projects are not without their risks. Participation is not always equal resulting in ‘free riding.’ There are also transaction costs (such as time, energy, meetings, and fraught negotiation of differences), blending of institutional or sector cultures, uneasy sharing of control, threats to credibility, and ethical principles (Toepell 2001). Moreover, communication difficulties (due to limited time, geographical distance, different experiences, frameworks, or assumptions) can lead to misunderstanding and offence.

The discussion now turns to examine how principles of communication and collaboration were applied in New Zealand, a small South Pacific nation (population

of 4 million), and how New Zealand responded to the ‘formative assessment movement’ in its particular context. Two examples of educational policies are examined, one in which the journey from policy conception to implementation was relatively smooth (development and implementation of National Curriculum Exemplars), and the other more problematic (National Standards).

## 6.4 Background to New Zealand Assessment Policy Formation

Since 1989, New Zealand has operated on a decentralised school system (Codd et al. 1995). Schools are individually governed by an elected Board of Trustees, with accountability to the taxpayer through the Ministry of Education’s stipulated National Administration Guidelines (NAGs) and National Education Goals (NEGs) (Ministry of Education 2013). Education policies are formulated centrally, with implementation details generally arising from sector-wide input and consultation. The examples used in this chapter illustrate this widespread consultation.

For example, an assessment document was developed by the Ministry after working with teachers, a range of curriculum and assessment experts, and other professionals throughout the sector:

a team of primary and secondary teachers prepared a draft...early drafts of the handbook were commented on by people administering Ministry contracts in teacher development in assessment, members of the Educational Assessment Secretariat Advisory Committee, and a number of people working in the field of assessment and evaluation. The chairs of the Minister’s Policy Advisory Groups for mathematics, science, and English, and other groups, were consulted in the final development phase. (Ministry of Education 1994, p. 4)

The document *Assessment Policy to Practice* (Ministry of Education 1994) was designed as a handbook to guide teachers and schools in developing their school-based formative assessment practices in line with requirements of *The New Zealand Curriculum Framework* (Ministry of Education 1993). A range of examples was given to illustrate how assessment principles might be translated into classroom practice. Amongst the examples given to foster formative assessment practice were exemplars. At that time, exemplars were little known or used in New Zealand.

Four years later (May 1998), a Green Paper, *Assessment for Success in Primary Schools*, was circulated to all New Zealand schools and educators beyond schools with an interest in assessment, as well as to parents, education organisations, and the general public. Feedback was collected (over a three month period) on proposals within the document. In the Green Paper the Ministry of Education proposed, amongst other initiatives, to develop exemplars of student work. The document stipulated the following purpose for exemplars: ‘To provide examples of criteria for assessing student work and to help teachers to decide whether the judgements they are making about student achievement are consistent with national expectations’ (Ministry of Education 1998, p. 20).

National expectations were not stipulated as such (hence the Ministry of Education's use of lower case letters in referring to national expectations) but implied was some connection with The New Zealand Curriculum (NZC). Whilst not specifically stated in these extracts, the exemplars were intended to accompany the five levels of NZC, in other words, from school Years 1–10 (Years 11–13 conform to the National Certificate of Educational Achievement examinations and associated requirements). Articulated, however, were principles governing use of exemplars.

Principles: Teachers need information to help them to identify whether their judgements about achievement are consistent with national standards. Teachers use assessment information to help them to evaluate the effectiveness of their teaching and learning programmes. The types of national assessment exemplars proposed here are: examples of student work and of teachers' marking criteria; examples of assessment activities.

Each exemplar would be linked to a particular curriculum statement, the relevant achievement objectives, level, desired process, and learning outcomes.... These national exemplars can be used to 'benchmark' individual student performance by setting out student responses to particular assessment tasks that teachers can compare with student responses at a classroom level.... National exemplars of student work will provide clear examples of expected levels of student achievement in relation to achievement objectives. The exemplars could also illustrate a range of student performance, from high-achieving to low-achieving students, providing realistic expectations of what students could achieve. (Ministry of Education 1998, p. 21)

Explicit within the purposes are the links to curriculum statements, and indeed, specific learning objectives and outcomes, as well as formative use of the assessment information. The perceptive reader will note the subtle insertion of neoliberal, economic terminology (e.g., outcomes, benchmarks, performance, and national standards). This is the first mention of national standards (albeit, a soft reference to it by the use of lower case letters). Contradictions are inherent within the purposes, such as 'benchmarks' (implied national norms), 'realistic expectations of what students could achieve' (implied developmental and learning progression type information); but more critically, the implicit accountability of teachers, 'identify whether judgements are consistent with national standards,' 'evaluate effectiveness of teaching and learning programmes,' 'clear examples of expected levels of student achievement in relation to achievement objectives.' Such contradictions were evidence of emerging divided political views, and of the influence of the National (centre right) Government of the time whose focus was on achievement (rather than learning)—product rather than process (Hill 1999).

Around this time there was an international and national political climate (Codd et al. 1995) suggesting the need for National Standards, with national testing in particular (as had occurred in England), and this was promulgated by the then National Government. However, shortly afterwards when the National Government was voted out (late 1999), the Labour (centre left) Government, not wanting to be associated with national testing, switched direction. Key personnel (including academics and Ministry of Education officials) convinced the Minister of Education at the time, Honourable Trevor Mallard, to redirect the appropriated funds away from national testing to further enhancement of formative assessment and development of

formative (rather than summative) assessment tools. Exemplars, and Assessment Tools for Teaching and Learning (asTTle), subsequently refined as e-asTTle, an electronic online assessment tool, were developed in New Zealand to assess students' achievement and progress in reading, mathematics, writing, and in Māori medium, for primary to lower secondary school levels (<http://e-asttle.tki.org.nz/> Accessed 5/3/2015).

With proposals for lower stakes and more classroom-based assessment tools, as well as an incoming Labour government with whom the teacher workforce is generally more comfortable working, teachers and unions were willing to engage in development processes. Furthermore, the Honourable Trevor Mallard (Minister of Education) highly valued a participatory model, met frequently with Ministry of Education officials, and requested extensive consultation with teacher unions, teachers, academics, and wider interest groups.

### **6.4.1 National Curriculum Exemplar (NCE) Development**

As signalled earlier, the scene was somewhat set from the teacher handbook, *Assessment Policy to Practice* (Ministry of Education 1994), for collaboration and communication, so the development of National Curriculum Exemplars over a four-year span (2000–2004) was a period of significant cross-sector collaboration. A large national team of developers was contracted, with sub-teams in each curriculum area: English, Mathematics, Science, Technology, Arts, and later Social Studies. These areas were developed alongside Māori medium teams. These teams predominantly comprised teacher education personnel but also contained private educational consultants and a few academics. A small national research team was contracted to work alongside the development team in order to provide ongoing feedback to the development work. The development subject areas were each led by a National Director, coordinated by their respective Ministry of Education curriculum facilitators, and the overall project was led by the Ministry of Education Manager of Assessment. The research team leader reported directly to this Ministry of Education person, since the purpose of the accompanying national action research project was to continually inform and improve processes and understandings associated with National Curriculum Exemplars (Poskitt et al. 2002).

Further indicators of collaboration were the intra and inter exemplar team processes established to create and refine the national curriculum exemplars.

#### **6.4.1.1 Intra-Exemplar Team Collaboration**

Regular (2–3 monthly) national *hui* (meetings) created opportunities for discussions within curriculum teams (e.g. only the Science group members together for one session) and across teams (e.g. representatives from each curriculum team responsible for collecting photographs) as well as various sector representatives (e.g. primary



teacher union) and researchers. Periodically, additional Ministry of Education and Learning Media personnel were present, either to provide input and advice (such as assessment principles and pedagogical strategies as well as publication considerations like colour, font size, size of jpeg files). The *hui* agenda included time for Ministry of Education updates, team reports of progress, research feedback, discussion of emerging issues, problem solving, and action planning. Whilst there was an element of accountability (hearing the progress of other teams; reminders of timelines and Ministry of Education expectations), the climate was more one of collaboration and mutual support. Teams experiencing difficulties obtained more time and, where necessary, supplementary resources, allocated by the Ministry of Education.

Because there were few exemplars used in New Zealand at the time, it was a journey of discovery—trial and error—as the exemplars evolved in each curriculum area. Exemplar development was a truly iterative process. Debates occurred within the curriculum teams and at the national *hui* related to:

- performance level of the student samples (e.g., whether the exemplars were to be typical student work, examples enhanced by teacher intervention and interpretation, or aspirational examples);
- representation of student learning (e.g., one-off samples or the record of learning development throughout a unit of work);
- representation of the curriculum level (eventually only Level 1 in English was differentiated into L1i, L1ii, L1iii in acknowledgement of the rapid learning stages of children in their first year of schooling; the other levels were kept broad. But debates continued as to whether samples ought to be at beginning, middle, or upper end of a curriculum level);
- risks of de facto curriculum (exemplification might risk privileging certain aspects of the curriculum; how extensive ought to be the coverage of the curriculum by exemplars, and therefore how many exemplars per curriculum and per level ought to be produced);
- rigorous discussions about the content of the curriculum statements and the sequencing of the curriculum (debates about learning progressions, depth and breadth of concepts, which concepts were deemed to be critical for next stage learning);
- purposes of the exemplars (teaching or learning tools, the extent to which pedagogical and assessment guidance might be provided; how formative to be in the statements and examples);
- how to enhance teachers' formative practice through using exemplars; and
- potential impact on students and teachers (threatening if seen to be too aspirational, concerns about potential use to judge teachers, how might teachers be guided in their interpretations and judgements on student work).

In between the national *hui*, curriculum teams continued the process of creating exemplars. Team members were geographically spread and communicated largely by email, fax, and phone but also held periodic face-to-face curriculum team meetings at which there were usually several Ministry of Education personnel in

attendance (to provide motivation, guidance, and influence direction of the development). Each of the curriculum teams utilized their educational networks (as former teachers, current advisors, or educational consultants) to trial ideas and collect samples of student work in schools to create exemplars.

Schools choosing to work in association with development teams were generally derived from informal networks (e.g., schools were either in previous or concurrent professional development contracts with the advisors or educational consultants, or had a professional relationship with the teacher education personnel through initial teacher education) and helped to develop exemplars by working in partnership with the development team member. Often the collaboration was manifested in co-teaching, with the curriculum developer teaching alongside the classroom teacher, sometimes modelling pedagogical strategies for the teacher and subsequently co-assessing the samples of student work generated from the teaching episode(s). There were several benefits from the collaboration: informal coaching and professional learning for the teacher (in curriculum content knowledge, pedagogical strategies, and formative assessment), participation in national assessment tool development, and mutual benefits for the curriculum developer who gained practical insights into classroom realities and appropriate expectations for students at a particular level. Furthermore, involvement of teachers ensured the exemplar initiative was better received by the teaching profession through professional credibility and deeper understanding of the nature and purpose of the exemplars. Deep ownership of the process and the products of exemplars occurred.

#### **6.4.1.2 Collaboration Beyond the Exemplar Development and Research Team**

The primary teachers' union, New Zealand Educational Institute (NZEI), the University of Canterbury, and the Ministry of Education partnered to sponsor a series of National Assessment Regional Seminars (July and October 2000) to promote the use of formative assessment practice. Around this period that terminology associated with formative assessment transformed into the more teacher-friendly phrase of 'assessment for learning'. Sector-wide involvement was evident in the participation of a wide range of representatives.<sup>1</sup> Furthermore, a Ministry of Education On-line Resource Centre was under development in late

---

<sup>1</sup>Teachers, principals, advisors, educational consultants, researchers; Ministry of Education National Assessment Advisory Group (comprising university academics, the Association of Colleges of Education, New Zealand Council for Educational Research, National Education Monitoring Project leader, *Te Runanga Nui o Nga Kura Kaupapa Māori o Aotearoa* [the National Collective Body of Māori language immersion schools embodying the language, philosophy and knowledge of Māori culture], professional development providers, primary, intermediate, and secondary school teacher and principal union representation, Schools Trustees Association, Catholic Education Office, New Zealand Area Schools Association, and the Independent Schools Council, as well as pertinent Ministry of Education personnel (Agenda notes, Ministry of Education 2000).

2000, *Te Kete Ipurangi* (meaning ‘on-line basket of knowledge,’ abbreviated as *tki*) (tki.org.nz).

The network web was cast further with a national trial of the exemplars. A random sample of 500 schools was invited to participate in the national consultation, representing every fifth school on the national school database apart from *kura* (schools in which Māori philosophy and language is used), who were involved in a separate consultation on Māori medium exemplars. Participating schools selected their degree of involvement (from reading exemplars, through to discussion and implementation in the classroom) in order to provide feedback by questionnaire, teacher interview, or researcher observations in schools. Some of these schools were part of the Assess to Learn (AToL) professional development programme that focused on and supported schools in implementing formative assessment practices within the classroom and at school level (Poskitt 2014). The exemplar research data not only informed the final iteration of the National Curriculum Exemplars but also the future professional development needs of teachers in the implementation phase (Poskitt et al. 2002). Moreover, participation in the trial enhanced teacher familiarity with, and ownership of, the exemplars through informal and formal professional networks, as well as use of the exemplars in improving learning and teaching.

Development of National Curriculum Exemplars was almost flawless in the process from policy formation to the final release of the documents to schools. Provision of sufficient time and resources allowed extensive communication and collaboration to occur, enabling educators to contribute to the process and product of exemplars, such that the policy was adapted to better align with school needs and aspirations.

#### **6.4.1.3 Implementation of National Curriculum Exemplars**

Implementation in schools and classrooms was mixed, however. Teachers and schools that participated in the development phases of the National Curriculum Exemplars or were involved in the AtoL professional development contracts (Poskitt 2014) were enthusiastic and used the exemplars extensively. But other schools were less sure of the purpose or how to use exemplars in teaching, learning, or assessment (despite frequent communication *from* Curriculum Update publications by the Ministry of Education). A missing factor was two-way communication. The environment within and outside of schools was changing rapidly. Budget cuts and Ministry of Education reprioritisation constrained resources available for centrally funded professional development to accompany exemplars. Concurrently, the national curriculum was reviewed, and pervasive thought in the school sector associated the exemplars with the former rather than revised New Zealand Curriculum (Ministry of Education 2007). As well, exemplar relevance was questioned.

## 6.5 Changing Policy Environment

By 2008 the political climate changed. Travelling policies and global influences were increasing momentum towards greater school accountability and the political need for national student achievement data, which was spurred by the growing influence of OECD research publications, and PISA results in particular (May and Sturrock 2002; Telford 2010). Concerns were emerging about the ‘tail of achievement’ (disparate achievement results in PISA indicating lower achievement rates of indigenous students) in New Zealand (Ell and Grudoff 2013; McNaughton and Lai 2009). Economic recession was widespread internationally, and the Labour Government was replaced by a National (centre right) Government that campaigned on implementing ‘National Standards.’ Moreover, there was a political imperative to deliver ‘National Standards’ in 100 days (‘National marks first 100 days’ 2009).

### 6.5.1 *Emerging National Standards Policy*

The recently revised New Zealand Curriculum (Ministry of Education 2007) drew on a literacy programme and numeracy framework with associated research on learning progressions. Development of further learning progressions was thought to be a relatively quick task. Due to the 100 day political imperative, only small groups in the sector were consulted. Serendipitously, the International Conference on Assessment was held in New Zealand (March 2009) at a critical time to shape the thinking of the Minister of Education, Honourable Anne Tolley. Several highly esteemed national (e.g., Terry Crooks) and international researchers (e.g., Kari Smith) in educational assessment were invited to a private meeting with the Minister. Frank discussions were held. Whilst it was noted that National Tests enable high reliability and accountability in the system, the low validity of national tests and their consequential narrowing of teaching and of the curriculum were risks to avoid. This discussion created an environment to encourage National Standards which were to be based on broad descriptions and teachers’ professional judgments. There was an understanding of the need to invest in teacher moderation of assessments (Chamberlain, personal communication, March 4, 2015). Broad descriptions in The New Zealand Curriculum, the existing educational climate, and practice of assessment for learning, along with contemporary pedagogical practice, were levers steering the government away from National Testing towards National Standards. Nevertheless, the National Government of the time was focused on results and speed, delivering on their campaign promises, rather than investing in process.

### 6.5.1.1 Conflicting Views on National Standards

In the early stages, unions were pleasantly surprised at the early conceptions, broad ideas, and descriptors (Chamberlain, personal communication, March 4, 2015). However, the speed of development, the narrow and limited consultative base, as well as teacher union suspicion of a National Government, meant the policy became highly politicised (“‘Hired gun’ in crossfire’ 2011; Stuff News 2012; Thrupp 2014). Widespread angst and political slogans created a negative climate that was too risky to trial emerging learning progressions (e.g., New Zealand Principals’ Federation 2009; NZEI, n.d.). Schools were forced to comply with the policy and used National Standards only superficially. Most schools did not appreciate nor optimise the formative assessment potential, despite moderation and assessment for learning assessment guidelines provided on the Ministry of Education website, *Te Kete Ipurangi* (tki). Indeed, at one stage some school leaders, including union and New Zealand Principal Federation members, were requesting national tests as a potential solution to the problem of prolonged hours of work by teachers involved in judging multiple samples of student work and associated moderation meetings. There was minimal teacher appreciation at that time of the respect the government had for teachers to make professional judgments on student assessments because the implementation process was characterised by compliance, minimal professional development, and a compressed timeframe.

At the school level, reaction to National Standards and Overall Teacher Judgments was mixed, largely influenced by the views of the principal (Poskitt and Mitchell 2012; Thrupp and White 2013). Reaction ranged from outright resistance through compliance to a determination to optimise the opportunity for their school. Schools with well-established professional learning systems and assessment for learning practices were more willing to engage in moderating samples of student assessment (Poskitt and Mitchell 2012). Issues they grappled with included: exploring how many samples of student work were sufficient to gain assurance of their professional judgments; distinguishing between evidence provided in the samples and deeper knowledge the teacher had of the students’ work; over what time frame to make judgments; against which reference points judgments would be made (difficult with generic New Zealand Curriculum and National Standards that provided space for regional interpretation); and time required for moderation. Many primary teachers struggled to reconcile a philosophical dilemma in assessment. On the one hand they held a general philosophy of positively viewing what students ‘could do’, generally used an ipsative (personal progress) framework and focused on ongoing learning. On the other hand, they were being asked to switch to a standards-based framework which required them to report standard levels to parents. Particularly conflicting for teachers was the need to report that some students were ‘below standard.’ Primary schools had for years focused on assessment for learning, and many were not comfortable with standards-based approaches. Teachers expressed fears of accountability, which were fuelled by union talk and international trends of connections to performance pay. New Zealanders have

deeply held values of equality, so talk of performance pay created consternation for many teachers.

### **6.5.1.2 Value of Communication and Collective Endeavour**

Despite public perception of steely obstinacy by the Minister of Education (Honourable Hekia Parata), who had a reputation amongst teachers for not listening, her professional relationship with key Ministry of Education personnel led her to take advice from them, as well as the National Standards Advisory Group (comprised of invited representatives of the sector), and selected experts. National Standards were kept broad, and some investment was made to assist teachers with moderation processes (e.g., Moderation modules on the Ministry of Education website, school advisors, and educational consultants supporting teachers with moderation processes).

With the passage of time, embedding of National Standards in the school system, informal sharing of Overall Teacher Judgments and moderation practice amongst teachers, release of a few years of national data on National Standards, and another New Zealand government election in which the National Government (centre right) was returned to power, teacher resistance diminished. Although some schools begrudgingly report National Standards, there is a growing appreciation for rich nationwide data, emerging benefits of a serious focus on struggling learners, and attention to outcomes of the ‘under-served’ students in the system (Chamberlain, personal communication March 4, 2015). Concerns still linger about credibility of the Overall Teacher Judgments data, especially regarding issues relating to dependability, reliability, risks associated with potential labelling of students, and consequential detrimental effects on motivation and learning. Some system-level trends indicate improving (albeit still lagging) outcomes for Māori and Pasifika students (Education Counts 2015). Policy analysts can now draw on system-wide data to discern areas of the sector in which targeted professional development and resources can make a difference.

## **6.6 Discussion**

These two examples of New Zealand policy formation and implementation reflect outcomes from different communication and collaboration processes. In the National Curriculum Exemplars example, open communication and extensive collaboration were hallmarks of the process, from the formation of the intention to develop exemplars through to final product development. Time and resources allowed for widespread consultation throughout the education sector (within and beyond schools) on an assessment for learning tool that had potential to benefit teachers and learners. Responsiveness to national and local input resulted in the tool (and the policy) having potential to influence not only assessment but also

pedagogical practices. The development process and evolving exemplar product is an example of symbiotic ‘policy adaptation’ where the purpose and use of exemplars expanded for the mutual benefit of government and educators. Exemplars evolved from the initial intent to ‘guide teachers in consistent judgements of national expectations’ (refer to Sect. 6.4) to a focus on educational partnership and learning, and using assessment information to inform teachers and learners:

The purpose of an exemplar ‘signals important features of student work to watch for, collect information about, and act on to support growth in learning; provides students, teachers and parents with a basis for discussing important qualities, aspects or indicators of learning; provides reference points that will support teachers’ professional judgments about the quality of their students’ work.’ (Ministry of Education 2002, p. 1)

The intent, design, and process of National Curriculum Exemplars development were well aligned. Input was sought and gained from within and across the educational sector, resulting in ideal conditions for willing cooperation in co-constructing tools for use in assessment *and pedagogical practice* in schools. This is a case of ‘policy adaptation’ and what Poskitt (2014) termed ‘collaborative assessment literacies’ (p. 562), namely: interactive learning serving the needs of students, educators and the community.

In contrast, implementation of National Standards with Overall Teacher Judgments was an example of misalignment of design and process, a policy adaptation possibly viewed as antibiosis by some educators. A plurality of views within and beyond the education sector led to uncertainty and heated debates, creating conditions for fractious insecurity and resistance. National and international politics did little to mitigate publicly expressed views, despite successes of a related policy approach in nearby Queensland, Australia (Klenowski and Wyatt-Smith 2014). Nevertheless, belief by the Ministry of Education in the design and its potential formative assessment benefits at school and system level prevailed. Support for this belief came indirectly from the NSAG (National Standards Advisory Group), the *Directions for Assessment in New Zealand Report* (Absolum et al. 2009), and assessment researchers arguing the importance of valuing teacher judgments, keeping assessments, curriculum, and pedagogical practices highly valid and broadly focused (e.g., Crooks 2010). Teacher perception was gradually modified and resistance lessened because of several factors: the publication of analysed National Standards data (Education Counts 2015) that correlated with data sources from National Monitoring and PISA; the re-election of a National (centre right) government, allowing National Standards to be further embedded; and the perceptible achievement gains for lower achieving Māori and Pasifika students. Indeed, principals and schools are increasingly *using* National Standards data to monitor and inquire into their school and classroom practices, to make adjustments to programmes, and to direct greater attention to student learning and achievement (Ward and Henderson 2011; Ward and Thomas 2013). In this case of policy implementation, the design had to withstand the storms of critical debate and resistance to become embedded before some users could perceive benefits. Rather than the policy being adapted *by* users, the users adapted *to* the policy.

## 6.7 Conclusion

The complexity of policy adaptation is now synthesised. Effective collaboration is based on sharing common goals and language, openness to receiving and contributing to mutual learning, reciprocating respect, and a growing sense of trust and friendship (Elliott and Woloshyn 2001; Schutz et al. 2001). Collaboration generates benefits of professional stimulation, mutual valuing of personnel, process and product, as well as accomplishment of new endeavours. Many of these benefits were evident in the National Curriculum Exemplars processes and outcomes. But the risks of refraining from collaboration were apparent through the early stages of National Standards implementation with differing purposes and goals, communication difficulties (largely due to restrictions on time and resources), threats to control and credibility, and conflicting values, leading to misunderstandings and indignation. But over time these risks and threats diminished as communication occurred and awareness of shared purposes emerged.

Borrowing and travelling policies are not therefore a matter of simple transfer or adoption by another nation. Rather, successful policy implementation necessitates adaptation of both the policy and the users. In the process from policy formation to implementation both the policy and the users need to adapt if full implementation is to occur. It is argued that participation in ‘policy adaptation’ is a necessary prerequisite to full implementation. However, the timing, the nature (of negotiation, co-construction, modification, and legitimisation processes), and the appropriateness of the adaptation will be situated within the conditions of politics, economics, culture, and education sector responsiveness *at the time*.

National Curriculum Exemplars had a smooth passage from policy conception through to implementation because users were able to adapt the policy (as well as their beliefs and practices) throughout the process to implementation. In a period of supportive political and economic conditions, educators across the sector were given opportunities to discuss, influence, and adapt the policy, with educators feeling professionally valued, willing to implement assessment for learning approaches, and readily perceiving benefits to learners. In contrast, with National Standards considerable policy adaptation occurred at system level prior to release to users. For example, National Testing policies from other countries were considered by the government, and those policies were significantly adapted—after consultation with education experts and Ministry of Education staff. Assessment experts and Advisory Groups exhorted the Minister of Education to value teacher professional judgement and maintain broad links between National Standards and the New Zealand Curriculum to minimise risks of narrowing the curriculum. These adaptations were either misunderstood or unknown by many educators. Because of perceived lack of communication and collaboration, the National Standards policy endured a rough passage. Furthermore, the economic, educational, and political climate had changed to one of greater plurality and uncertainty. Greater levels of accountability and fewer opportunities to shape and adapt the policy undermined teachers’ perceptions of trust and professional value, resulting in prolonged



resistance and less authentic implementation. To the users there appeared to be a greater emphasis on the user having to adapt *to* the policy rather than the policy being adapted *by* them. Although these issues are currently less intense, perceptions of suspicion about the purpose and use of the National Standards data still persist (Thrupp 2014).

What can be learned from the New Zealand experience on policy adaptation? There are lessons about timing as well as levels of communication and collaboration, valuing professionals through participation in policy processes, and consequences for the extent of policy implementation. Policy implementation in every country is challenging, but the challenges can be diminished when time and resources are invested in communication and collaboration, particularly in the early stages of policy formation and implementation. Without it, implementation takes longer, coercive approaches become necessary (e.g., legislation requiring schools to use and report National Standards), residual resentment remains in some quarters, and there is less openness to perceiving benefits to learners. Clarity of purpose is pivotal, along with opportunities for participation of various parties with interest and jurisdiction (namely teachers, educational consultants and advisors, academics, unions, and policymakers; with potential to involve parents and students). Willingness to inform and negotiate across these layers not only develops shared language, understanding and respect but also forges collective commitment for the betterment of student learning. When the design is clearly communicated, and the inclination for *mutual adaptation of policy and user* is present, the storms of contradiction and conflict can be overcome resulting in widespread implementation and use of formative assessment policies in schools.

## References

- Absolum, M. (2010). *Clarity in the classroom: Using formative assessment for building learning-focused relationships*. Auckland, New Zealand: Hachette New Zealand Ltd.
- Absolum, M., Flockton, L., Hattie, J., Hipkins, R., & Reid, I. (2009). *Directions for assessment in New Zealand: Developing students' assessment capabilities*. <http://hdlvoked.edu.au/10707/183949>. Accessed March 12, 2015.
- Adie, L. (2014). The development of shared understandings of assessment policy: Travelling between global and local contexts. *Journal of Education Policy*, 29(4), 532–545. doi:10.1080/02680939.2013.853101.
- Bridges, D. (2014). The ethics and politics of the international transfer of educational policy and practice. *Ethics and Education*, 9(1), 84–96. doi:10.1080/17449642.2014.890274.
- Clarke, J. (2014). A principle to govern educational policy: Abolishing National Standards. *New Zealand Principal*, 29(1), 29–32.
- Codd, J., McAlpine, D., & Poskitt, J. (1995). Assessment policies in New Zealand: Educational reform or political agenda? In R. Peddie & B. Tuck (Eds.), *Setting the standards: Issues in assessment for national qualifications* (pp. 32–58). Palmerston North, New Zealand: The Dunmore Press Ltd.
- Crooks, T. (2010). Educational assessment in New Zealand schools. *Assessment in Education: Principles, Policy & Practice*, 9(2), 237–253. doi:10.1080/0969594022000001959

- Divala, J. (2014). The ethics and politics in educational policy development and practice. *Ethics and Education*, 9(1), 97–101. doi:10.1080/17449642.2014.890276
- Department of Education. (1989). *Assessment for better learning: A public discussion document*. Wellington: Department of Education.
- Education Counts. (2015). *National Standards Achievement Results 2014*. <http://www.educationcounts.govt.nz/statistics/schooling/national-standards>. Accessed March 12, 2015.
- Ell, F., & Grudoff, L. (2013). The politics of responsibility: Teacher education and ‘persistent underachievement’ in New Zealand. *The Educational Forum*, 77(1), 73–86.
- Elliott, A., & Woloshyn, V. (2001). Collaboration: A vehicle for professional development. In M. Richards, A. Elliott, V. Woloshyn, & C. Mitchell (Eds.), *Collaboration uncovered: The forgotten, the assumed and the unexamined in collaborative education* (pp. 177–189). Westport, CT: Bergin & Garvey.
- Froese-Germain, B. (2010). *The OECD, PISA and the impacts on educational policy*. Canadian Teachers’ Federation (NJ1). [https://scholar.google.co.nz/scholar?q=The+OECD%2C+PISA+and+the+Impacts+on+Educational+Policy&btnG=&hl=en&as\\_sdt=0%2C5](https://scholar.google.co.nz/scholar?q=The+OECD%2C+PISA+and+the+Impacts+on+Educational+Policy&btnG=&hl=en&as_sdt=0%2C5). Accessed November 6, 2015.
- Hill, M. (1999). Assessment in self-managing schools: Primary teachers balancing learning and accountability demands in the 1990s. *New Zealand Journal of Educational Studies*, 34(1), 176–190.
- ‘Hired gun’ in Crossfire. (2011, April 7). [Editorial]. *Waikato Times*. <http://www.stuff.co.nz/waikato-times/opinion/4858816/Editorial-Hired-gun-in-crossfire>. Accessed March 5, 2015.
- Klenowski, V., & Wyatt-Smith, C. (2014). *Assessment for education: Standards, judgement and moderation*. London: SAGE.
- McNaughton, S., & Lai, M. (2009). A model of school change for culturally and linguistically diverse students in New Zealand: A summary and evidence from systematic replication. *Teaching Education*, 20(1), 55–75.
- May, S., & Sturrock, F. (2002). *PISA 2000: The New Zealand context*. Wellington, New Zealand: Ministry of Education Research Division.
- Ministry of Education. (1993). *The New Zealand curriculum framework*. Wellington, New Zealand: Learning Media.
- Ministry of Education. (1994). *Assessment policy to practice*. Wellington, New Zealand: Learning Media.
- Ministry of Education. (1998). *Assessment for success in primary schools: Green paper*. Wellington, New Zealand: Government Print.
- Ministry of Education. (2002). *The national exemplars trial*. Wellington, New Zealand: Learning Media Ltd.
- Ministry of Education. (2007). *The New Zealand curriculum*. Wellington, New Zealand: Learning Media.
- Ministry of Education (2013). *National administration guidelines and national education guidelines*. <http://www.minedu.govt.nz/theMinistry/EducationInNewZealand/EducationLegislation/TheNationalEducationGuidelinesNEGs.aspx>. Accessed March 12, 2015.
- National marks first 100 days in office. (2009, February 26). *New Zealand Herald*. [http://www.nzherald.co.nz/business/news/article.cfm?c\\_id=3&objectid=10558844](http://www.nzherald.co.nz/business/news/article.cfm?c_id=3&objectid=10558844). Accessed March 5, 2015.
- New Zealand Educational Institute (NZEI). (n.d.). <http://www.standupforkids.org.nz/primary-school-league-tables/>. Accessed March 5, 2015.
- New Zealand Principals’ Federation. (2009). <http://www.nzpf.ac.nz/national-standards>. Accessed March 5, 2015.
- Ozga, J., & Jones, R. (2006). Travelling and embedded policy: The case of knowledge transfer. *Journal of Educational Policy*, 21, 1–17. doi:10.1080/02680930500391462
- Poskitt, J. (2014). Transforming professional learning and practice in assessment for learning. *The Curriculum Journal*, 25(4), 542–566. doi:10.1080/09585176.2014.981557

- Poskitt, J., Anthony, G., Brown, M., & Taylor, K. (2002). *Milestone report to the Ministry of Education: National consultation on draft national exemplars*. Palmerston North, New Zealand: Massey University.
- Poskitt, J., & Mitchell, K. (2012). New Zealand teachers' overall teacher judgements (OTJs): Equivocal or unequivocal? *Assessment Matters* 4, 53–75.
- Schutz, A., Murray, N., Woloshyn, V., Elliott, A., Morgan, N., & Haskins, B. (2001). Not just smooth sailing: Issues in collaboration. In M. Richards, A. Elliott, V. Woloshyn, & C. Mitchell (Eds.), *Collaboration uncovered: The forgotten, the assumed and the unexamined in collaborative education* (pp. 219–235). Westport, CT: Bergin & Garvey.
- Steiner-Khamsi, G. (2012). Understanding policy borrowing and lending: Building comparative policy studies. In G. Steiner-Khamsi, & F. Waldow, F. (Eds.), *World yearbook of education 2012: Policy borrowing and lending in education* (pp. 3–17). New York: Routledge.
- Steiner-Khamsi, G., & Waldow, F. (Eds.). (2012). *World yearbook of education 2012: Policy borrowing and lending in education*. New York: Routledge.
- Stuff News. (2012, September 22). How New Zealand students rate. *Stuff News*. <http://www.stuff.co.nz/national/education/7715044/How-New-Zealand-schools-rate>. Accessed March 5, 2015.
- Telford, M. (2010). *PISA 2006: Scientific literacy—How ready are our 15 year olds for tomorrow's world?*. Wellington, New Zealand: Ministry of Education Research Division.
- Thrupp, M. (2014). At the eye of the storm: Researching schools and their communities enacting National Standards. *New Zealand Journal of Educational Studies*, 49(1), 6–20.
- Thrupp, M., & White, M. (2013). *Research, analysis and insight into National Standards (RAINS) project. Final report: National Standards and the damage done*. Wellington, New Zealand: NZ-EI. <http://www.waikato.ac.nz/wmier/research/projects/the-rains-research-on-national-standards>. Accessed April 7, 2016.
- Toepell, A. (2001). Cooperative versus collaborative research: Assessing the difference. In M. Richards, A. Elliott, V. Woloshyn, & C. Mitchell (Eds.), *Collaboration uncovered: the forgotten, the assumed and the unexamined in collaborative education* (pp. 61–80). Westport, CT: Bergin & Garvey.
- Waldow, F. (2012). Standardisation and legitimacy: Two central concepts in research on educational borrowing and lending. In G. Steiner-Khamsi & F. Waldow (Eds.), *World yearbook of education 2012. Policy borrowing and lending in education* (pp. 411–423). New York: Routledge.
- Ward, L., & Henderson, A. (2011). *An evaluation of network learning communities—Technical Report*. Wellington, New Zealand: Ministry of Education. <http://www.educationcounts.govt.nz/publications/curriculum/an-evaluation-of-network-learning-communities-technical-report>. Accessed March 12, 2015.
- Ward, J., & Thomas, G. (2013). *National Standards: School sample monitoring and evaluation project*. Dunedin, New Zealand: Maths Technology Limited. [http://www.educationcounts.govt.nz/publications/series/National\\_Standards](http://www.educationcounts.govt.nz/publications/series/National_Standards). Accessed March 12, 2015.

# Chapter 7

## More Than Good Intentions: Policy and Assessment for Learning in Scotland

Ernest Spencer and Louise Hayward

**Abstract** The major challenge in Scotland's long history of well-intentioned policy has always been implementation, in particular the realisation of a constructive and effective relationship across research, policy, and practice. Scottish experience provides a basis for radical changes, of potential international significance, in assessment policies to ensure better practical orientation to learning. The chapter considers critically the relationship between assessment policy rhetoric in the Curriculum for Excellence (for students aged 3–15) and provision of practical guidance and professional learning opportunities. It draws on understanding of what matters in the process of change gained from previous Scottish experience in the Assessment is for Learning programme. Evidence from a study of early Curriculum for Excellence assessment practice, Assessment at Transition, shows how the design, findings, and conduct of that project have led to some collaborative action by researchers, policymakers, and practitioners to make effective implementation of key assessment policy intentions more likely, despite the inadequacy of the support originally provided. The argument then moves beyond steps to help the implementation of current Scottish policy by proposing a number of major changes to the purposes and content of typical 'traditional' assessment policies and practices not only in Scotland but in many countries.

### 7.1 Implementing Change in the Assessment Is for Learning (AifL) Programme

From 2002, an Assessment is for Learning (AifL) programme was developed and implemented across the Scottish Education system. The approach strongly promoted the idea that the crucial purpose of *all* assessment of individual learners, formative and summative, and of *all* evaluation of educational provision (e.g., in school self-evaluation, in analysis of external examination results, and in inspections) was

---

E. Spencer (✉) · L. Hayward  
University of Glasgow, Glasgow, UK  
e-mail: ernest.spencer@glasgow.ac.uk

essentially to provide evidence about and to contribute to the improvement of learning (Scottish Government 2002, Circular 02). One key aim was the development across the system of assessment for formative purposes during ongoing classwork. This was integrated in a cyclic process with planning, learning/teaching, and identifying next steps for individual learners and for groups/classes. Clear learning aims and success criteria, feedback, reflection, and self- and peer assessment were strongly promoted as contributing to learning to learn, to engagement, and to self-confidence. This broad ‘formative assessment’ aspect came to be called in practice assessment for learning (AfL) and incorporated the idea of assessment as learning, involving self- and peer-assessment activities through which students come to develop reflection and independence as learners.

The programme was sophisticated. It aimed to enable and support teachers as they developed their own professional understanding of effective assessment integral to the process of learning (Hayward 2007). It drew on research about assessment, certainly, such as that reported in Black and Wiliam (1998a, b), Black et al. (2002, 2003), The Assessment Reform Group (2002), Harlen and Deakin Crick (2002, 2003). It was also informed by emerging research on transformational change. Senge and Scharmer (2001), in a meta-analysis of public and private organisations which were perceived to have transformed their practices, emphasized three key ideas that should underpin organisational change:

- a shared set of guiding principles across all participating communities—for the AifL programme, that meant shared principles across research, policy, and practice communities
- collaborative projects that relate to people’s professional lives, offering real contexts for participants to deepen their understanding of principles and ideas and of what these might mean for their practice
- all parts of the system working together—for example, in the Scottish system, the guiding principles of assessment for learning should be endorsed and promoted by national and local government and by Her Majesty’s Inspectors of Education (HMIE) during school inspections (Hutchinson and Hayward 2005), as well as by head teachers and teachers.

Consistent with the research findings on transformational change, the development was led by Scottish Government staff with significant expertise in assessment and sought to build innovative, supportive communities with clear practical tasks. Each Local Authority identified clusters of schools (primary and secondary) where staff discussed how they were going to explore the issue and produced a plan for what they intended to do. Scottish Government gave a small grant to each school to use as it wished to support its activities. Most used it to buy time for teachers to work together, to read the research evidence, to talk through ideas, and to learn from one another’s practice. Teachers involved were also invited to network meetings, which offered opportunities to discuss ideas with others from across Scotland and from schools in England where similar approaches had been or were being developed.

The support network for teachers was bolstered by other networks: one of local authority co-coordinators who each had responsibility for the assessment for learning activities in their own area; one of development officers, working for the national curriculum body, Learning and Teaching Scotland, who offered practical support to schools and local authorities; and a higher education research and development network, with representatives from across Scotland (including the authors of this chapter), whose tasks included ensuring that the programme was informed by research evidence and that initial teacher education and wider professional development programmes took account of the assessment for learning development. The possible risks to the success of assessment for learning were themselves the subject of research while the programme was being implemented. The Scottish Government-appointed Managing Committee of the programme (of which Louise Hayward was a member) commissioned research from the Higher Education network and from consultants, who gathered evidence about particular aspects of the development. These research activities provided feedback to all the various participants in the programme as it was being taken forward. There were parents' and students' networks, meeting in open forums across the country. Quality assurance personnel (HMIE and local authority officers) were actively involved in promoting and evaluating what was happening. At national level there were meetings with the Association of Directors of Education in Scotland (ADES) and a network of civil servants to promote internal consistency and continuity in national policies.

### ***7.1.1 Assessment for Learning Success***

Complex as the support structure for the AifL development was, it seemed to work. The overlapping networks ensured that no one group had total responsibility for supporting another and teachers had several groups to whom they could look for support. Good evidence emerged that the programme did enable many teachers to develop pedagogy incorporating assessment for learning and that many learners benefited. The evaluations carried out by the Institute of Education, University of London (Hallam et al. 2004), and the University of Strathclyde (Condie et al. 2005) were very positive. Teachers found that being involved in the development was professionally rewarding and that it made very positive differences to students' commitment to learning and the quality of their work. They were delighted to be able to focus on what mattered, learning and teaching. A common theme in their reports was that, in working through what its principles meant for them in planning and leading work in their classrooms, they had developed a deep understanding of how to use assessment to support children's learning (Hayward et al. 2005). The assessment for learning initiative had such an impact that it was described in the press as 'a quiet revolution in Scottish Education' (Henderson 2005).

Evidence is limited about the extent to which effective assessment for learning became and continues to be deeply embedded in classrooms across the country after

the end of the formal development period. As Curriculum for Excellence was being developed and implemented, the national inspection system (now conducted by Education Scotland) has focused principally on evaluating curriculum development and the effectiveness of schools' internal self-evaluation and improvement processes. Inspection reports have provided little information about the nature and quality of assessment. However, some evidence is available from a small number of research activities. These include:

- An evaluation of Strategies for Early Arithmetical Learning (SEAL) used in years 1–4 of a small sample of primary schools in a local authority area. The East Lothian Council SEAL report (Hayward et al. 2014) described the development of young students' self-awareness as learners through stimulation of thinking, teachers' response to it, and student collaboration in the early stages of primary education. It identified a range of *interacting* factors in the learning and teaching experiences as crucial to success—not any one 'magic bullet.'
- The Highland Project, a Scottish Government-funded study in primary schools (Hayward 2012) which highlighted the significance in effective assessment for learning of 'student voice' and 'the learner's perspective.'
- A study (Hayward et al. 2009) commissioned by the Scottish Qualifications Authority (SQA) which described teachers in several Highland Council secondary schools enabling students to make very successful formative use of the formal published criteria for Intermediate or Standard Grade examinations at age 16 and Higher examinations at age 17 in a range of subjects.
- The Assessment at Transition (AaT) project (Hayward et al. 2012) commissioned by the Scottish Government in the early stages of the practical implementation of Curriculum for Excellence. The focus here was assessment to provide evidence about learning and progression at the point when students pass from primary education into secondary, at age 11–12. Both teachers and students interviewed made frequent references to assessment for learning principles and activities.

Though these four research studies involved assessment for learning activities which varied to some extent according to the age of the students and the subject contexts, there were evident common factors for success. There was a cycle of learning/teaching which incorporated collaborative enquiry by the learners and integrated assessment for learning as a constantly occurring activity, ensuring secure, shared understanding before moving on to next steps. The students were active, motivated learners, involved in collaboration with the teacher in co-regulation of learning and in clarification/agreement of aims and criteria. They explained their own thinking, engaged regularly in teacher-student dialogue and self- and peer assessment, and gave the teacher feedback about their own learning experiences which could be beneficial to the future learning of others.

Hayward and Spencer (2014), drawing on Black and Wiliam (2009), have argued that it is possible—and desirable—to think of assessment for learning as essentially consisting of three recurring generic activities: stimulating learners to

think about the topic, the curricular learning they are pursuing; finding out what and how they are thinking (often through dialogue); and identifying with them next steps for new challenges, more effective thinking, fuller, more certain grasp of what is being learned. Assessment for learning thus contributes significantly to psychological development and independent learning: the quality of interactive feedback and reflection is a critical feature in learning activity and assessment for learning develops the orientations, abilities, and confidence characteristic of independent and collaborative learning. This conception of assessment for learning is very much in keeping with what Marshall and Drummond (2006) called the ‘spirit of assessment for learning,’ as opposed to the mere use of ‘techniques’ such as ‘traffic lighting.’ It is heartening that the four research studies mentioned do indicate that, at least in localised contexts, assessment for learning matching the broad characteristics set out by Hayward and Spencer (2014) and by Marshall and Drummond (2006) has been going on across the whole range of education, from early primary school to the senior phase of secondary.

There was, however, another kind of important finding from Assessment at Transition. The students interviewed were asked which change in current assessment practice they most wanted to see; almost all said they wanted more one-to-one conversation with their teacher about their learning. This answer suggests that there may be less of the essential dialogue actually happening than one might have thought from the many references to assessment for learning in the interviews. Even during the development of assessment for learning in the original programme from 2002, which was evaluated as generally successful (Hallam et al. 2004; Condie et al. 2005), there were indications that some teachers tended to interpret assessment for learning as simplistic use of certain ‘techniques,’ such as ‘traffic lighting’ or ‘thumbs up/down,’ while others engaged students in much more sophisticated thinking and dialogue about their learning (Hayward et al. 2005). Overall, the Curriculum for Excellence practical reality is probably complex: a continuum from deep understanding of assessment for learning, curriculum and pedagogy, and the role of self-and peer assessment and agency in these, to superficial use of techniques and/or, as suggested by the students interviewed in the AaT project, provision of too little or unhelpful feedback to learners.

### ***7.1.2 Successful Implementation of Change***

The AifL development provided a rich context for deepening understanding of critical factors in successful change programmes. Hayward et al. (2005) reported on a study involving interviews with teachers, head teachers, and local authority coordinators to identify the characteristics of the programme which had facilitated its success, and Hayward and Spencer (2010) drew on this report to reflect and comment on important generic factors that contribute to successful change programmes. The central ideas emerging from this commentary relate to three concepts of integrity, all of which are crucial to successful action:



**Educational Integrity** Teachers and head teachers believed that involvement in assessment for learning had led them to a much sharper focus on learning and learners than on teaching and getting through the curriculum. They reported shifts in power relationships in their classrooms, with learners having far greater responsibility. They recognised that they themselves had developed greater concern for understanding what and how students were thinking and building from there. They saw these changes as manifestly valuable for students' education.

**Personal and Professional Integrity** Teachers had been attracted by the methodology of assessment for learning as professionals. They felt that their views mattered as it moved forward, that they were consulted as an essential part of the process. They welcomed opportunities to talk through problems and ways of doing things, both with teaching colleagues from their own school and elsewhere and with assessment experts supporting the development. Although many admitted to initial concerns about losing control, they enjoyed the more relaxed atmosphere in their classrooms and their own learning—some expressed regret that they had only now come to realise the effectiveness of assessment for learning after many years' experience in the classroom. Many referred to the challenge of what they were trying to do but spoke very enthusiastically of their enjoyment at seeing students learn more effectively.

**Systemic Integrity** It was clear that systemic integrity was important to the commitment of all those who took part in the AifL development. Head teachers and teachers involved reported explicitly that knowing their local authority, HMIE (the inspectors), and government were supportive of the programme gave them confidence to change.

Hayward and Spencer (2010) also argue that the complexity of the process of change is actually itself a desirable characteristic and that attempts to make it 'manageable' often in effect oversimplify it and damage and reduce the richness of its effects. They therefore advise against 'traditional' development models such as 'pilot and roll out' or (unsupported) 'cascade.' Things that matter in the process of change matter to *all* participants in *every* phase of the development: it is therefore essential that the kinds of support, collaboration, and professional learning opportunities which benefit those participating in the early stages of a new development should be sustained and built into the normal professional life of all those who take part at later stages.

Before leaving this account of the assessment for learning development it is worth noting that both evaluations of it (Hallam et al. 2004; Condie et al. 2005), as well as identifying positive features, highlighted factors with potential to constrain its success. The potentially negative factors remain problematic ten years later in the context of Curriculum for Excellence. A particular challenge was the relationship between formative assessment and the demands of summative assessment. Practical concerns were raised about finding time for teachers to deepen their professional understanding of the assessment system, to allow them to engage with principles and ideas, and to work collaboratively to explore how best to put these into practice. Provision of time was perceived to be related to the availability of funding for teacher cover, and this too was identified as a concern. Finally, both evaluations

highlighted the potential danger of bureaucracy dominating learning and teaching and problems emerging from the polarisation of assessment purposes in secondary schools, most evident in the senior years, where high stakes external examinations were perceived to dominate practice.

## 7.2 Assessment in Curriculum for Excellence

The Curriculum for Excellence seeks to provide a coherent curriculum from ages 3–18 (Scottish Government 2008a, b, c, d, 2009). 3–18 experience is divided into two phases, Broad General Education to age 15 and the Senior Phase, involving pursuit of qualifications through assessments provided and/or verified and accredited by the Scottish Qualifications Authority. This chapter addresses only issues relating to the 3–15 phase.

The development and implementation of Curriculum for Excellence began formally with the Education Minister's acceptance of initial proposals for principles and purposes (Scottish Executive 2004) and schools began to implement it in 2009. Critical questions arise about the extent to which the whole process has conformed to the 'Integrity' model described by Hayward and Spencer and, in particular, the extent to which the preexisting improved assessment for learning has been incorporated into Curriculum for Excellence practice.

### 7.2.1 *Curriculum for Excellence Policy*

Curriculum for Excellence policy aims to promote outcomes essentially similar to those of the earlier AifL programme—breadth, challenge, and application in learning, increased emphasis on process and skills, rather than memorisation of content, and teacher professionalism. Key purposes are to encourage students to become 'Successful Learners,' 'Confident Individuals,' 'Effective Contributors,' and 'Responsible Citizens.'

Curriculum progression for young people 3–15 is described in Experiences and Outcomes in eight curricular areas across five levels, early, first, second, third, and fourth (which offers possibilities for choice for those who have completed the Experiences and Outcomes at the third level). Experiences and Outcomes also exist at the various levels for three cross-curricular areas, literacy, numeracy, and health and wellbeing 'across learning' (Scottish Government 2009). The levels descriptors are not highly specific about objectives: their deliberate broadness is partly intended to encourage teachers to develop their professionalism in deciding how to pursue and achieve the curriculum aims in various ways. The level descriptors are also intended to be not strictly age related: some Experiences and Outcomes have the same wording across two or three levels.

From the beginning policy documentation (Scottish Executive 2004) emphasized individuals' needs and entitlements and warned against unnecessary bureaucracy, tight specification of curricular outcomes and standards, and overly complex systems. The Building the Curriculum (BtC) series of documents, 1–4, published by the Scottish Government in 2008 (covering curriculum areas, active learning in the early years, learning and teaching, and skills for learning, life, and work) consistently referred to the need to ensure that both curriculum and assessment focused on learning and were in alignment. *BtC5: A Framework for Assessment* (Scottish Government 2011) and its three additional supporting documents argue the same case. They also specify that teachers should make summative assessments of students' attainment of the Curriculum for Excellence levels; that they should report this attainment to parents/guardians; and that students should be enabled to develop their own profiles of successful learning at Primary 7 (age 11) and Secondary 3 (age 14): these personal profiles are intended to record achievements, whether within school or elsewhere, that students themselves value highly. Policy also emphasizes strongly the importance of moderation of the quality of assessment activities and judgements.

### ***7.2.2 Assessment Issues in Curriculum for Excellence***

The Assessment at Transition (AaT) Report (Hayward et al. 2012) suggested that there were major challenges in putting key aspects of policy into practice in the early stages of the implementation of the Curriculum for Excellence. A range of issues relating to assessment of learning, described below, may well have been, and may still be, constraining teachers' and learners' opportunities to engage in valuable assessment for learning activities because they are very time consuming. They may also have been leading to erroneous evaluations of both the progress of individual learners and the overall quality of achievement in a class or school because they suggest significant weaknesses in the quality of assessment of learning which was taking place.

There was evidence that teachers needed support to be able to think of all the factors contributing to very effective learning as part of one coherent process. These factors include curriculum planning, design of learning tasks, agreement on success criteria, making judgements about whether young people's work meets the criteria, helping learners to reflect on their own learning and to identify next steps and, on occasion, summarising success and progress, and moderating judgements about these.

In some local authorities, the demand for frequent overall summative (level) judgements—three or four times per year, with the ostensibly good intention of monitoring individual students' progress—dominated assessment activities and actually militated against teachers' developing professional understanding of curriculum, pedagogy, and assessment as a coherent whole. A few local authorities had divided each level into three sublevels (Developing, Consolidating, Secure) and

required progress records using these subcategories. Teachers were consistently very uncomfortable with the requirement to make levels judgements, with or without the sublevels. They argued that they had no confidence in their own understanding of standards or of appropriate summative assessment processes. They could find in the national documentation no clear definition or exemplification of standards and no helpful guidance on how to proceed to make a level judgement or to record progress in other ways. Some tried to use the level as a kind of grade, evaluating individual pieces of work as, e.g., 1st or 2nd Level, rather than making a 'best fit' judgement about achievement of the level based on a body of evidence.

Despite the absence of national testing from the Curriculum for Excellence 3–15 Phase, almost all local authorities perceive a continuing 'requirement' for accountability to local politicians in terms of 'hard evidence': hence the demand on teachers to provide regular summative assessment of levels achievement, even though there were wide variations and probably little validity in the methods used. The perceived need for 'hard data' has led to widespread use by local authorities (including those also gathering teachers' levels judgements information) of standardised tests to provide 'accountability evidence.' Standardised test results are also often used to 'track' students' progress. These tests are regarded as appropriate for these purposes despite the fact that they typically test only aspects of literacy and numeracy and are not designed to assess the specified outcomes of Curriculum for Excellence.

One other factor emerged as significant from the Assessment at Transition discussions with staff in secondary schools. Although at that point the new National Qualifications arrangements for the Senior Phase (15–18) had not yet been published, it was clear that many secondary teachers were awaiting sight of these arrangements before deciding firmly how to structure and teach the curriculum and assess progress in the first three years of secondary education (12–15). National Qualifications exert a powerful influence on learning/teaching in secondary schools.

### **7.3 Curriculum (and Assessment) for Excellence and the Process of Change**

A major challenge for Curriculum for Excellence was to merge the new ideas about curriculum and learning processes with the preexisting successful assessment for learning practice. A member of the Board responsible for the design and implementation of the Curriculum for Excellence Programme has suggested that the Board seems to have *assumed* there would be a natural integration of effective assessment for learning, rather than actually planning for it (Hayward 2015). Another kind of policy mistake (made in a different forum) led to a public declaration that the Curriculum for Excellence Experiences and Outcomes (Es and Os) represent *assessment* outcomes and constitute the standards statements for the various attainment levels. In fact, the writers of the Es and Os did not intend this use

of them—they were meant as curricular guidance only—and, in many cases, they make poor standards statements (for example, some have identical wording across two or even three levels). The policy documents for assessment were not developed in conjunction with the curricular thinking and the publication of the curriculum guidance: they emerged later, almost as an afterthought. This delay might be interpreted as a strategy designed to ensure that teachers thought about the curriculum before thinking about assessment. It could also be argued, however, that it separated thinking about curriculum and assessment. The assessment guidance was typically couched in general exhortatory terms about the desirability of good assessment rather than providing specific guidance on steps to achieve it. The absence of advice on *how* to achieve what the policy advocated was particularly noticeable in relation to the processes of assessment of learning—deciding on and recording learners’ success in achieving the standards (or levels) of work expected.

In these circumstances it would seem that we cannot say that all three types of ‘Integrity’ elaborated by Hayward and Spencer (2010) are fully apparent in the implementation of Curriculum for Excellence and of assessment within it. A case can be made that many aspects of the Curriculum for Excellence initiative and its development are indeed manifestly educationally valuable; and it is true that the programme explicitly seeks to encourage and promote teachers’ individual professionalism and their teamwork. However, interviews with teachers in the Assessment at Transition project showed that in general they did not feel that the implementation of the programme sought to engage them professionally in the way that some previous developments, including AifL, had done. Many felt that they were being required to make assessment of learning judgements about level attainment without appropriate professional guidance and support and that they were simply being told to use their professionalism without the opportunity to develop it appropriately. They felt that the kind of support they needed to develop successfully the professionalism the Curriculum for Excellence explicitly aims to promote had been misjudged. They did not have a sense that the whole system—policymakers, the two agencies with assessment responsibilities, inspectors, local authorities, head teachers—was in fact working effectively together to make assessment in Curriculum for Excellence highly successful.

#### **7.4 Current Action: Research, Policy, and Practice in Collaboration**

The AaT project was designed to improve the alignment of research, policy, and practice *during implementation of the new curriculum*: it aimed (1) to find ways of helping schools to implement policy and (2) to use research (both a comprehensive literature review, which underpinned the project, and findings on schools’ and local authorities’ practice) to inform and influence desirable policy changes through interaction with Scottish Government and local policymakers in seminars.

Members of the research team have been interacting with national and local policymakers since the completion of the project. Both policy and practice communities involved in the study emphasized the importance of a *limited number of sharply focussed* action proposals. The four key areas where action has been stimulated are:

1. Developing teacher professionalism in bringing together curriculum and assessment
2. Managing learning and progression at transition
3. Building trust in professional judgement
4. Ensuring intelligent accountability in Curriculum for Excellence.

The researchers argued that, to relieve pressure on teachers and students of too much summative assessment and create space for effective assessment for learning, levels judgements should be *infrequent* (three times in the Broad General Education phase, 3–15). Drawing on a wide range of research evidence, including Morrison et al. (1994) and MacPhail and Halbert (2010), they proposed that the most valid means of determining level achievement was to use a ‘best fit’ approach, considering whether a body of classwork matched the description of key curricular learning and the quality of work required, which would be described and exemplified. Social moderation arrangements for discussion of judgements were also proposed. The AaT report also highlighted some key principles for accountability arrangements. It emphasized the need to focus more on the quality of educational experiences and less on test results, in order to ensure consistency with Curriculum for Excellence aspirations and to avoid negative washback on classroom activities.

The project findings were directly addressed by Education Scotland’s (2013) publication for schools and teachers *Assessing Progress and Achievement of Levels in the 3–15 Broad General Education*. A national initiative, established to support the development of policy, used research evidence on assessment analysed in the AaT literature review to offer guidance on using professional judgement to make decisions about achievement of levels. Work from Australia (Wyatt-Smith et al. 2010; Colbert et al. 2012) and New Zealand (Crooks et al. 2009) was particularly influential. The initiative identified significant aspects of learning (SALs) in curricular areas, developed, through working groups of teachers and subject experts and progression frameworks (rubrics) for these SALs. It has begun to bring together annotated exemplification of student work to illustrate attainment of a curricular level in terms of these rubrics. This was a major recommendation of the AaT research in response to teachers’ very explicit requests for such support. Further, drawing on evidence from both the review of literature and from practice which had been developing in the schools and local authorities, the initiative proposed a process of professional learning through learning communities. In the first phase of this development of professional learning, groups of teachers have been meeting to discuss the exemplification being produced at national level, feed back their own views on it, and enhance their abilities to judge accurately the achievement of a level. A second phase has followed in which teachers bring evidence from their classrooms to discuss their own professional judgements with colleagues.

In addition, the research team has interacted with key local authority staff around the country, inviting them to reflect on the validity, usefulness, and advantages or disadvantages of current arrangements in their areas for gathering assessment information from schools.

The ‘direction of travel’ represented by this action in the period 2012–2015 is helpful. Significant adjustments are being made to policy and implementation processes as a result of research commissioned by the policymakers, which was designed to provide useful feedback to the Curriculum for Excellence programme as it was put into practice in the varying and dynamic contexts of local authorities and schools. Sustained support will be necessary to enable these adjustments to have the desired effects.

## **7.5 More Radical Steps?**

The Scottish experience raises some more significant issues than just how to make the best of an existing flawed local system. We propose for international consideration potentially radical changes to traditional ways of thinking about and conducting assessment in primary and early secondary education.

### ***7.5.1 Curriculum, Pedagogy, and Assessment as a Coherent Whole***

The unsatisfactory nature of (at least) the processes by which Scottish teachers are making summative judgements about learners’ progression and overall achievements reinforces the importance of taking forward thinking about curricular aims and about assessment simultaneously. To enable learning, teachers need clear ideas about progression routes, understanding of effective pedagogical steps to stimulate learners’ thinking and action, and familiarity with means of gathering evidence about their learning and acting on it to promote further learning. An intention to develop teachers’ professionalism in these areas requires provision of significant time to enable them to interact collaboratively. This need has, of course, a significant financial implication in terms of staff numbers. Guidance is needed about action to take to ensure the coherence of curriculum, pedagogy, and assessment, as is mediation of this guidance (by people with appropriate expertise) such that teachers’ professional thinking and their interactions are stimulated and can influence national thinking, without just imposing a wholly centralised, top-down system. This kind of approach can work effectively only when researchers, policymakers, and practitioners really do come together in a synergy that optimises their various strengths. It seems clear that in the development or modification of curricula there is need for such integrated action to optimise pedagogy and assessment as well as learning aims.

### 7.5.2 *Focus on Assessment for Formative Purposes*

Perhaps the most crucial question is this: is there just too much demand for summative assessment of learning in primary and early secondary education?

We have indicated above the local pressures on teachers to provide summative information frequently. We have also shown something of the complex process necessary if teachers are to make summative assessments validly and dependably, involving the application of a ‘best fit’ model to a portfolio of work and a good deal of discussion with colleagues in moderation meetings. It is common practice in many countries for teachers to write reports on every student’s progress for parents/guardians and for the information of the next teacher at least once a year, which may indicate an overall grade or level and provide brief comment and an orientation to next steps. In some administrations (as in Scotland) there may be in addition a need to help students prepare profiles of their personal achievements or to contribute to the reporting process themselves. If the essential purpose of assessment is to promote rich learning and to enable both students and teachers to build effectively on prior learning (at transition from primary to secondary education or transition from one class to another), we should be asking ourselves whether we really need summative assessment of levels performance—or standardised test results—in primary and early secondary education. Accountability arrangements could be refocused on the quality of learning rather than test results, or reshaped so that information about systemic performance comes solely from national monitoring surveys.

### 7.5.3 *Prioritising Assessment Activities*

Taking account of the complexity and time-consuming nature of summative assessment processes and of reporting, the AaT research team raised the question whether it is actually *feasible* for primary school teachers to conduct valid and reliable assessment of levels achievement or to write detailed descriptions of progress across the whole range of curriculum areas. There are major time-consuming assessment activities that serve little purpose and should be ended. Recognising that time is inevitably limited, the team argued that it is important to *prioritise* assessment activities, a process that entails stopping doing some things in order to make it possible to do other, more desirable, things well.

A particularly significant finding of the project gives support to the idea that some quite radical prioritisation could enable teachers to focus all their assessment activities very directly on helping students to develop as learners. Secondary teachers told the researchers that they did not use detailed reports sent by primary teachers to help them decide what and how to teach new students. They said they used any indication they received of the level a student had achieved only to give them a very rough idea of her/his current abilities—for example, to place her/him in



a high or low set for mathematics. What they did find very useful as a basis for building on students' prior learning was (a) information about previous curriculum coverage; (b) one-to-one or small group discussions with the students about what they had learned before they reached the secondary school; and (c) good professional interaction with their primary teacher colleagues about curriculum planning and ways of teaching. Interestingly, these views of the teachers harmonised well with the students' main concern, that they wanted more dialogue with their teachers, and with research highlighting the importance of discussion with students at the time of transition to secondary school, such as that reported in Doddington et al. (1999), and Demetriou et al. (2000).

There is a strong case for simplifying policy messages on assessment of progress and achievements in primary and early secondary education in the following ways.

### **7.5.3.1 Focus on the Learner**

- Keep the focus sharply on the learner and on ensuring her/his progression.
- Continue to develop assessment for learning and learner independence strongly.
- Promote the idea of reporting to parents only through *discussions* based on manageable annotated portfolios of student work, with very broad categories of comment about overall progress—this would require time and organisation, but significant time currently devoted to report writing would be saved.

### **7.5.3.2 Change Expectations at Transitions**

In order to be able to build on students' prior experience, 'receiving' teachers should

- Have clear curriculum coverage information.
- Discuss previous work with students—e.g., focusing on manageable folios of work.
- Engage actively in a professional learning community with colleagues, including those in the 'other' sector (primary/secondary).

### **7.5.3.3 No Overall Grades or Levels**

- Abandon attempts to make overall grade or level judgements about individual students.
- Ensure the curricular progression pathways are clearly defined and assess students' success in achieving key learning specified within them.

### 7.5.3.4 National Monitoring of System-Wide Standards of Achievement

- Develop well-designed national monitoring assessments to be administered on a sampling basis to provide information about system-wide standards of achievement.
- Such arrangements can be designed to ensure that, over an agreed period of time, all schools participate and receive individual feedback (as in Finland).

Radical changes of this type could be successfully introduced only with very full discussion with all stakeholders in the system. There would need to be a process of engagement with school managers, teachers, students, and parents to demonstrate the gains which could ensue in terms of dialogue about learning. There would also be a need to negotiate a very different use of teachers' time over an academic year, involving more meetings with parents and students in place of time committed to report writing.

We believe that changes such as those we have outlined would benefit Scottish education. We invite readers elsewhere to consider whether comparable prioritisation and streamlining of their country's use of assessment would similarly benefit students.

The relationship amongst research, policy, and practice is complex. It is all too easy for well-intentioned policy to result in practices that are very different from original aspirations. Bringing policy and practice into closer alignment will take more than good intentions. Research projects, such as those cited in this chapter, have a role to play in that process. Using research to explore the interrelationship of policy and practice as an evidence base to inform future action can help to realign policy aspirations and practice in schools and classrooms. Action based on evidence is the only way to build education systems that are truly learning systems.

## References

- Assessment Reform Group. (2002). *Assessment for learning: 10 principles*. <http://www.aaia.org.uk/content/uploads/2010/06/Assessment-for-Learning-10-principles.pdf>. Accessed July 30, 2015.
- Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice*, 5(1), 7–73.
- Black, P., & Wiliam, D. (1998b). *Inside the black box: Raising standards through classroom assessment*. London: GL Assessment.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2002). *Working inside the black box: Assessment for learning in the classroom*. London: King's College London School of Education.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Buckingham: Open University Press.

- Colbert, P., Wyatt-Smith, C., & Klenowski, V. (2012). A systems level approach to building sustainable assessment cultures: Moderation, quality task design and dependability of judgment. *Policy Futures in Education*, 10(4), 387–402.
- Condie, R., Livingston, K., & Seagraves, L. (2005). *The assessment is for learning programme: An evaluation*. Edinburgh: Scottish Executive Education Department. <http://www.scotland.gov.uk/Publications/2005/12/0792641/26439>. Accessed June 9, 2014.
- Crooks, T., Darr, C., Gilmore, A., Hall, C., Hattie, J., Smith, J., & Smith, L. (2009). *Towards defining, assessing and reporting against national standards for literacy and numeracy in New Zealand*. Canterbury: The New Zealand Assessment Academy, University of Canterbury.
- Demetriou, H., Goalen, P., & Rudduck, J. (2000). Academic performance, transfer, transition and friendship: Listening to the student voice. *International Journal of Educational Research*, 33, 425–441 (Special Issue: School Transitions and Transfers).
- Doddington, C., Flutter, J., & Rudduck, J. (1999). Exploring and explaining ‘dips’ in motivation and performance in primary and secondary schooling. *Research in Education*, 61, 29–38.
- Education Scotland. (2013). *Assessing progress and achievement of levels in the 3–15 broad general education*. Glasgow: Education Scotland. [http://www.educationscotland.gov.uk/Images/CfEbriefing2\\_tcm4-730387.pdf](http://www.educationscotland.gov.uk/Images/CfEbriefing2_tcm4-730387.pdf). Accessed July 30, 2015.
- Hallam, S., Kirton, A., Pfeffers, J., Robertson, P., & Stobart, G. (2004). *Interim report of the evaluation of programme 1 of the Assessment is for Learning Development programme: Support for professional practice in formative assessment*. London: Institute of Education, University of London. <http://www.scotland.gov.uk/Publications/2004/10/19947/42989> Accessed July 30, 2015.
- Harlen, W., & Deakin Crick, R. (2002). *A systematic review of the impact of summative assessment and tests on students’ motivation for learning*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London, In Research evidence in education library.
- Harlen, W., & Deakin Crick, R. (2003). Testing and motivation for learning. *Assessment in Education: Principles, Policy and Practice*, 10(2), 169–207.
- Hayward, E. L. (2007). Curriculum, pedagogies and assessment in Scotland: The quest for social justice ‘Ah kent yir faither’. *Assessment in Education: Principles, Policy and Practice*, 14(2), 251–268. doi:10.1080/09695940701480178
- Hayward, L. (2012). Assessment and learning: The learner’s perspective. In J. Gardner (Ed.), *Assessment and learning* (2nd ed.). London: SAGE Publications.
- Hayward, L. (2015). Assessment is learning: The preposition vanishes. *Assessment in Education: Principles, Policy and Practice*, 22(1), 27–43. doi:10.1080/0969594X.2014.984656
- Hayward, L., Boyd, B., MacBride, G., & Spencer, E. (2009). *Just making them think: A tension between teaching and assessment in the high stakes stages*. Glasgow: Scottish Qualifications Authority (Research Report 13). [http://www.sqa.org.uk/sqa/files\\_ccc/Report%2013\\_Make\\_Them\\_Think.pdf](http://www.sqa.org.uk/sqa/files_ccc/Report%2013_Make_Them_Think.pdf). Accessed July 30, 2015.
- Hayward, L., Hutchinson, C., & Spencer, E. (2014). *Evaluation of stages for early arithmetical learning (SEAL) for East Lothian Council* (Research Report). Glasgow: University of Glasgow.
- Hayward, L., Menter, I., Baumfield, V., Daugherty, R., Akhtar, N., Doyle, L., et al. (2012). *Assessment at transition*. Glasgow: University of Glasgow.
- Hayward, L., Simpson, M., & Spencer, E. (2005). *Assessment is for learning: Exploring programme success*. Edinburgh: Scottish Executive Education Department. <http://wayback.archive-it.org/1961/20100806040920>, [http://www.ltsotland.org.uk/Images/AifL%20Exploring%20Programme%20Success\\_tcm4-354151.pdf](http://www.ltsotland.org.uk/Images/AifL%20Exploring%20Programme%20Success_tcm4-354151.pdf). Accessed July 30, 2015.
- Hayward, L., & Spencer, E. (2010). The complexities of change: Formative assessment in Scotland. *Curriculum Journal*, 21(2), 161–177. doi:10.1080/09585176.2010.480827
- Hayward, L., & Spencer, E. (2014). Assessment for learning. In M. Carroll & M. McCulloch (Eds.), *Understanding learning and teaching in primary education*. London: SAGE Publications.
- Henderson, D. (2005, December 9). New tests still to prove their worth. *Times Educational Supplement Scotland*. <https://www.tes.com/article.aspx?storycode=2170216>. Accessed October 19, 2015.

- Hutchinson, C., & Hayward, L. (2005). The journey so far: Assessment for learning in Scotland. *Curriculum Journal*, 16(2), 225–248. doi:10.1080/09585170500136184.
- MacPhail, A., & Halbert, J. (2010). We had to do intelligent thinking during recent PE: Students' and teachers' experiences of assessment for learning in post-primary physical education. *Assessment in Education: Principles, Policy and Practice*, 17(1), 23–39. doi:10.1080/09695940903565412.
- Marshall, B., & Drummond, M. J. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education*, 21(2), 133–149. doi:10.1080/02671520600615638.
- Morrison, H. G., Busch, J. C., & D'Arcy, J. (1994). Setting reliable national curriculum standards: A guide to the Angoff procedure. *Assessment in Education: Principles, Policy and Practice*, 1(2), 181–199. doi:10.1080/0969594940010204.
- Scottish Executive. (2004). *A curriculum for excellence: Ministerial response*. Edinburgh: Scottish Executive.
- Scottish Government. (2002). Circular no. 02: Assessment and reporting 3–14. <http://www.scotland.gov.uk/Publications/2005/06/2393450/34518>. Accessed July 30, 2015.
- Scottish Government. (2008a). *Building the curriculum 1: The contribution of curriculum areas*. Edinburgh: Scottish Government.
- Scottish Government. (2008b). *Building the curriculum 2: Active learning in the early years*. Edinburgh: Scottish Government.
- Scottish Government. (2008c). *Building the curriculum 3: A framework for learning and teaching*. Edinburgh: Scottish Government.
- Scottish Government. (2008d). *Building the curriculum 4: Skills for learning, skills for life and skills for work*. Edinburgh: Scottish Government.
- Scottish Government. (2009). *Assessment for curriculum for excellence—Strategic vision, key principles*. Edinburgh: Scottish Government.
- Scottish Government. (2011). *Building the curriculum 5: A framework for assessment* (2nd ed.). Edinburgh: Scottish Government.
- Senge, P., & Scharmer, O. (2001). Community action research: Learning as a community of practitioners. In P. Reason & H. Bradbury (Eds.), *Handbook of action research: Participative inquiry and practice*. London: SAGE.
- Wyatt-Smith, C., Klenowski, V., & Gunn, S. (2010). The centrality of teachers' judgement practice in assessment: A study of standards in moderation. *Assessment in Education: Principles, Policy and Practice*, 17, 59–75. doi:10.1080/09695940903565610.

**Part II**  
**Professional Development**  
**and Collaborative Learning**  
**About Assessment**

# Chapter 8

## Building Capacity: Professional Development and Collaborative Learning About Assessment

Dany Laveault

**Abstract** This chapter presents the topic addressed in Part II of the book entitled *Professional Development and Collaborative Learning about Assessment*. It begins by documenting why assessment for learning is challenging and why professional development is important. It then moves on to define what skills are needed to use assessment information to support learning and what practices are suitable indicators of such a competence. A co-regulation model is used to characterize different variations of professional development and collaborative learning. An entire section presents research evidence on the factors and conditions that enable or facilitate successful professional development and collaborative learning. As a conclusion, new perspectives from different chapters of Part II are provided, and recommendations are made on how to move forward in this domain.

### 8.1 Introduction

Professional development (PD) is a major component of policy enactment. It plays such an essential role in meeting the challenges of AfL implementation that it requires a part of this book of its own. Policy enactment and PD are closely interrelated, and there are several means by which PD may help enact AfL policy.

Although certain basic learning and training on classroom assessment occur at the preservice teacher training level, Part II is purposefully devoted to in-service teachers, as much of the learning on AfL can barely be achieved during the teacher education years and will primarily need to be supported and reinforced over many ensuing years. PD also involves all stakeholders, principals, head teachers, supervisors, and researchers who learn from each other during a PD activity. For instance, a school principal attending PD activities may develop an awareness of what AfL professional development consists of in terms of learning challenges and

---

D. Laveault (✉)  
University of Ottawa, Ottawa, Canada  
e-mail: dany.laveault@uottawa.ca

an understanding of the difficulties teachers will encounter in implementing new assessment strategies.

Part II is also about collaborative learning because, as this part of the book will show, it frequently occurs as a component of a global PD strategy. Furthermore, collaborative learning appears to be an essential condition for sustained and durable learning of AfL. Collaborative learning provides opportunities for policy enactors and stakeholders to exchange feedback information during and after professional development.

## 8.2 Why Assessment for Learning Professional Development Is Important and Challenging

Professional development is central to meet the challenge of implementing educational policies and causing the change that is needed to make classroom assessment play an important role in supporting student learning. It is also required for some of the reasons previously noted in the introduction on policy implementation (Chap. 2) and for the following reasons:

- *The complexity of AfL competence.* AfL necessitates the integration and mobilization of a series of skills in several domains of teaching practice: classroom management, instruction and learning, subject matter knowledge, curriculum, program of studies, and inevitably, assessment. This statement implies that AfL is a highly complex multi-faceted competence that requires much time to develop and that requires a form of accompaniment (mutual support, mentoring, and collaboration among peers) from a variety of stakeholders.
- *The lack of teachers' basic knowledge of assessment practice.* AfL PD requires prerequisite knowledge on classroom assessment. Several researchers have reported that the basic knowledge on which AfL could be developed is seriously lacking among in-service teachers. For instance, Schneider and Meyers (2012) reported that less than one-third of the teachers they sampled showed the skill to properly align the learning task to the stated learning goals. Yap (2007 in Schneider and Meyers 2012, p. 3) also observed that '34 % of teachers in their study could not accurately interpret a state standard of their own choosing.' Consequently, proper AfL PD may only occur once teachers have assimilated certain general prerequisites on classroom assessment; otherwise, the chances of successfully implementing AfL are seriously hampered.
- *The cognitive complexity of teachers.* Teachers' capacity to use AfL successfully depends on teachers' communication skills and their ability to interact with students. Reynolds (1970) has shown that there is a direct link between teachers' cognitive complexity and their verbal interactions with students. The implications of his study are that 'classes taught by cognitively simple instructors contain a high percentage of lecture, drill, teacher-direct talk and allow for

relatively little acceptance and use of student ideas, little student-initiated talk...’ (Reynolds 1970, p. 63).

- *The development of expertise.* Direct experience in using multiple sources of assessment information is a necessary condition for AfL learning. It takes time to develop *experts’ schemata* that will ‘allow them [teachers] to weight information so that its saliency and utility are determined quite quickly. In teaching, such skill in processing information is necessary because of the complex, dynamic, information-rich world of the classroom’ (Carter et al. 1987, p. 156).
- *Impact of PD on AfL.* AfL, as a part of a mandatory policy, is relatively new. Consequently, our knowledge base of what are the best in-service conditions to assist teachers in learning and using AfL consistently is relatively new and emerging. Moreover, ‘student-achievement related research is sparse and has not supported strong causal conclusions regarding the effect of teacher professional development in formative assessment practices on student achievement’ (Schneider and Meyer 2012, p. 3). This statement means that much progress still needs to be made in teacher PD before we can capitalize on the full impact of AfL on students’ learning outcomes.

### 8.3 Teachers’ Skills and Professional Development Challenges

The capacity to use assessment information appears to be a real challenge for teachers when the purpose is to support learning. For instance, Heritage et al. (2009) show that although teachers may agree on a student’s learning problem, they may not concur regarding what is the best next step in his or her learning progression. According to Herman et al. (2010), teachers may lack the knowledge base or skills they need to reach the proper decision on learning progression. In both studies, however, teachers were required to make decisions based on assessment tools that they did not help design. The results may have been different if the same teachers had been involved in the design of the assessment tasks. A recent OECD report (Looney 2011, p. 29) concludes that ‘teachers need to develop skills not only to identify individual student learning needs, but also to respond to them’ (p. 29).

Teachers surveyed in an OECD study (Looney 2011) noted the importance of being more systematic in their approach to classroom assessment because the most effective interactions with students are the result of careful planning. Poor task design or testing with no clear objectives does not allow teachers to collect the information they need to improve their decisions regarding how they can adapt their teaching to support learning. Furthermore, lack of awareness of a task’s cognitive demand may make it nearly impossible for teachers to add domain-specific information on students’ systematic errors in a manner in which it can be reported with adequate levels of reliability and validity. According to Webb and Jones (2009), ‘facilitating change in teachers’ assessment practice is not so much a resource



problem as it is a problem of... helping teachers develop a “designers’ eye” for selecting, adapting and designing tasks to assess student understanding’ (p. 3). Although such a capacity will improve teachers’ assessment practice in AfL, it is equally important that teachers experience ‘varied views on student work over time and in different contexts’ so they can ‘identify patterns in thinking and problem solving’ (Looney 2011, p. 9). Teachers need to assess and explore a range of potential causes to develop an appropriate teaching intervention.

Both Webb and Jones (2009) and Shepard (2006) focus our attention on what should be the impact of PD on teachers’ assessment practices. At this point, an ‘impact model,’ which describes what is expected from teacher professional learning (PL), is a necessary condition for appropriate PD goal setting, PD feedback and monitoring of teachers’ assessment practices. Hill (2011) used Davies and Busick’s description of teachers’ best AfL practices as such a model.

In this model, teachers:

- begin with the learning goals in mind;
- engage students in the process of understanding the learning destination, considering the evidence of learning towards those goals and considering what quality work looks like;
- directly involve their students in co-constructing criteria, self-assessing in relation to the criteria, giving themselves and others specific descriptive feedback, applying feedback to improve their work, collecting evidence or proof of their ongoing learning, and summarizing what they have learned and presenting it to peers, parents, and the community;
- use the assessment information they and their students gather to make informed teaching decisions that will engage all students and help them learn (Hill 2011, p. 353).

Whether or not one agrees with a specific impact model, it is important that such models be made explicit not only to teachers but also to other stakeholders for the purpose of studying, comparing, and assessing the professional learning outcomes of different PD strategies. It is nearly impossible for teachers, as well as for other stakeholders, to judge whether certain methodologies of PD have been successful if achievement targets have not been clarified and made explicit from the beginning.

## 8.4 A Regulation Model of Professional Development

In the same way that the concept of regulation provides a relevant framework to address the issues of how to make the best use of assessment information to support student learning, a regulation model may be used to define how teacher PD can support teachers’ learning and move AfL practices forward. Consequently, within a regulation framework, teacher PD must consider the following three components of learning regulation:

- *Goals.* Well-defined professional learning goals are needed to set realistic targets and to consider teachers' prerequisite knowledge and skills. This means that the PD goals will target different skills and mastery levels depending on teachers' existing levels of professional learning and zone of proximal development. However, such goals must set the targets of change to the levels of standards that attest to valid professional achievement in AFL. That is why standards such as those found in the *National Board of Certification* (United States) or as in Hill's impact model (Hill 2011) are so essential.
- *Self-awareness, self-monitoring, and sources of feedback for teachers' PD.* Teacher PD requires that a form of progress monitoring be in place to remain on target and to make the necessary adjustments as a function of emergent needs and new challenges. Fundamental to this process is the requirement that teachers become self-aware of their own assessment practices and develop a capacity to self-monitor their teaching and assessment practices; therefore, they can focus on the means by which to improve. Through a spiral of mutual influences, self-awareness will lead to modifications in the teachers' self-belief systems such as *internal causal attribution, perception of self-efficacy, and learning orientation*, which, in turn, will contribute to increased self-awareness. Different feedback sources (e.g., students, colleagues, and parents) can also assist teachers in adjusting their self-perceptions and to have a better idea of their own progression towards the professional learning goals.
- *Action, remediation, and teacher agency.* Professional learning (PL) will occur, or will occur to different extents, depending on the action-decision taken by teachers to change their practice and move toward the preset targets. Teachers may be allowed more or less discretion or agency on their PD. For instance, PD may be targeted to very specific school priorities with minimal consideration of a teacher's learning needs or capacity to adjust to changing conditions. Conversely, goals may also be set by the teacher with the school principal and in accordance with a formal assessment of teaching and a well-established career plan.

A regulation model of PD can describe different ways that teacher professional learning may unfold with self-, co- and shared regulation. Butler and Schnellert (2012) 'found this model particularly useful in characterizing how teachers might engage in iterative cycles of knowledge generation, through which they coordinate tacit and more explicit forms of knowledge' (p. 1207).

There are instances where regulation does not apply or is moderately relevant. PD that targets the short-term transmission of knowledge content with no specific expectations of teachers' learning is one such case. These instances also occur whenever certain training sessions are not followed up to monitor the degree of implementation of training goals. One can barely talk of regulation of learning when there are no specific or well-understood learning targets or criteria to determine what the professional learning expectations are and what will be needed to determine that they have been achieved.

## 8.5 Variations in Teacher Professional Development and Collaborative Learning

Teacher PD may be influenced by a series of factors: the nature of learning expectations, teachers' intrinsic or extrinsic motivation or the extent that the goals are clearly understood. Teacher PD also varies in accordance with the opportunities afforded for teachers' self-reflection and collaborative learning and as a function of teachers' agency.

Depending on the prevailing educational context, successful PD may take a variety of forms and occur in a variety of situations, many of which will be unforeseen. For instance, Brookhart et al. (2010) have reported that, contrary to their expectations, professional learning occurred in a highly scripted environment:

The fact that they [teachers] were able to find ways to modify their instruction, even when faced with a highly scripted context, in order to respond to student learning needs, points to a critical aspect of professional learning – change of belief. (p. 53)

One other such instance of a highly scripted environment occurs under the conditions that prevail at the U.S. *National Board Certification* (NBC). Sato et al. (2008) reported that teachers' voluntary participation in the certification program resulted in sustained changes in classroom assessment practices closely associated with AfL. Being required to self-report on their actual assessment practices in relation to the Board's standards of practice helped teachers develop an increased awareness of how they assess students and helped them become more self-critical. Compared with a control group, such changes were maintained for at least a year after the end of the certification process:

The teachers who experienced the National Board Certification process reported that the requirements of analyzing their classroom practice with a focus on assessment as defined by the National Board teaching standards introduced them to new ways of viewing the role that assessment plays in their everyday instructional interactions. The process of videotaping their teaching and analyzing it also brought elements of their practice into sharper focus. (Sato et al. 2008, p. 698)

While increased self-awareness of NBC teachers' assessment practice appears to be an important factor of change in AfL, the NBC professional standards also contributed to establish well-defined goals of professional achievement. Ingvarson (1998) suggests that professional standards 'provide goals for professional development that constitute a stable, challenging, and long-term agenda for professional development' (p. 130).

The use of artefacts to scaffold teacher PD has also been reported to be a valuable starting point. However, professional learning must be supported by more profound changes in classroom assessment:

Traffic lights and the thumb tool were both used for self-assessment. Superficially these mediating artefacts appear to be easy to introduce but unless a culture of honesty and openness about learning has been developed they actually put pressure on students and can have a negative effect. (Webb and Jones 2009, p. 180)

Providing teachers with mediating artefacts should also take into consideration the degree of teacher expertise in using AfL with their students. The artefacts may have a positive impact on novice teachers as a first step in learning about AfL; however, an overreliance on tools may actually become counterproductive:

A focus on the tools of formative assessment rather than its philosophy, so that the use of tools such as traffic lights and peer assessment became fixed as the object, emerged as a barrier to development. Thus the choice of mediating artefacts and their order of introduction was very important for enabling formative assessment. (Webb and Jones 2009, p. 178)

Mediating artefacts to engage both students and teachers in the change process may prove useful at the beginning but may be limiting in the long term if more challenging goals and a change in the classroom environment are not progressively introduced. Even if these options provide teachers with a form of short-term results, such easy fixes may short-circuit the need for deeper reflection and inquiry on AfL. This approach may explain why PD may lead to superficial incarnation of AfL and to teachers' conformance to the letter of AfL without a true grasp of its spirit (Earl and Timperley 2014). Whatever their level of expertise or PD, it appears that teachers appreciate a progressive approach to PL: 'teachers stressed the importance of not being prescriptive about starting points and sequences of development' (Webb and Jones 2009, p. 179).

Mediating artefacts are not the sole means to scaffold teachers' professional learning. Collaboration with peers not only helps teachers to develop 'within school coherence in terms of teaching and learning' (Parr and Timperley 2010, p. 160) but also allows teachers to benefit from their peers' experience and knowledge. Collaborative learning triggers co-regulation processes, which Butler and Schnellert (2012) define as follows:

Co-regulation occurs when a social agent provides support to or "scaffolds" another's engagement in cycles of inquiry, whether as an equal partner or as a mentor. From this perspective, it could be argued that working within a network or community of inquiry creates conditions for teachers not only to access rich resources, but also to engage together in developing practice and learning. (p. 1208)

AfL PD may involve two different forms and degrees of teacher collaboration.

1. *Learning from others* involves low to moderate levels of collaborative learning. Learning occurs as a result of in-service teacher training or mandatory PD sessions such as required courses to update skills. Teachers interact primarily with an expert, a researcher, or a mentor. Such training models may also involve a certain degree of interaction among peers as part of the training or as teamwork during or in between training sessions.
2. *Learning with others* involves moderate to high levels of collaborative learning. In its simplest form, learning with others may consist of a dyad between peers of equal status for team teaching or a form of social moderation of students' assessment. In its most elaborate forms, learning with others may involve a whole community of teachers working together as professional learners in the

same school and on the same subject matter or any topic of common interest of their own choosing. Collaborative inquiry into student learning illustrates one such model (Brookhart et al. 2010).

In practice, collaboration among teachers, leaders, and other professionals needs time and facilitating conditions. Flexibility and a mix of different opportunities for collaboration are likely what work best:

development of collaborative relationships within a networked structure is definitely not automatic. Teachers required time, space, and opportunities to work with colleagues and leaders within and across schools. Also important was the opportunity to seek out colleagues with similar levels of commitment and/or complementary knowledge. (Butler and Schnellert 2012, p. 1215)

## 8.6 Enablers of Assessment for Learning Implementation

Regardless of the circumstances of PD, changes in teacher self-confidence, as in the Brookhart et al. (2010) study, and working towards clearly defined and challenging goals, as in the Sato et al. (2008) study, helped teachers to become more self-aware and self-critical of their assessment practice and focus on what needed to be improved.

Although research evidence on PD thus far shows that professional learning occurs in a variety of conditions, including unexpected ones, it also reveals that certain common factors or conditions are at play to enable or prevent it. A ‘cascade’ model of school-based assessment change where ‘teachers receive training as facilitators and then act as in-school facilitators and work with other teachers on a school-wide basis’ (Hill 2011, p. 349) leads to superficial changes. According to Hill (2011), a facilitation model ‘tailored to meet the needs of the school’ appears to provide the best conditions for collaborative learning: ‘change is more effective if facilitators start where teachers are at in terms of their assessment practices, and work from there through collegial inquiry’ (p. 349). Such beginning conditions are indeed necessary in the regulation of PD to ascertain that teachers and facilitators agree on the PD goals, the need for change, and the criteria that will be used to ensure that PD goals have been achieved.

The regulation of PD involves helping teachers become more self-aware of the gap between their assessment practice and the PD goals and providing them with useful feedback on their PL progression towards these goals. Timperley et al. (in Hill 2011) identified seven facilitative contexts for teacher PD that are necessary but not sufficient on their own: ‘external expertise; being engaged in learning rather than volunteering to change; challenging prevailing discourses (of learning and teaching); participating in a professional community of practice; alignment with trends in wider policy and research; and active school leadership’ (p. 348). Combined, all of these contexts will make it easier for teachers to become

self-aware of their teaching practices, challenge their prevailing views and pre-conceptions, and obtain targeted feedback and support.

In Hill's study (2011), school participants' views of enablers of AfL implementation included an important role for the school leadership such as the involvement of the senior staff management team and the role of the principal as the 'conductor' of change. Teachers believed that the school principal needed to be 'assessment-literate and well familiar with assessment for learning practices' (p. 356). Indeed, the school leadership can hardly change teachers' preconceptions on assessment and set appropriate PD targets unless school leaders can convey a clear representation of what AfL really means and involves. Without AfL literacy, they can hardly help teachers become more self-aware or provide them with the necessary challenges and useful feedback they need.

## **8.7 New Perspectives on Professional Development and Collaborative Learning About AfL**

As our knowledge base on AfL progresses, PD becomes crucial in ensuring that AfL is delivered to students to its fullest extent. Thus, the more we learn regarding PD of assessment for learning, the more we realize that AfL is a complex competence that will require a lot of time, support, and collaborative work among teachers, school leaders, and other professionals. Meeting the challenges of AfL implementation can hardly be accomplished with a few expert conferences, sporadic training sessions, or workshops. This statement appears to be a conclusion with which all contributors of Part II would easily agree.

One of the major lessons learned from Part II contributions is likely the value of alternating PD seminars and classroom experiences to allow teachers to develop new conceptions of assessment. This rotation of theoretical-practical-reflexive learning opportunities appears to be a major feature of several contributions in Part II (Chap. 9—DeLuca et al., Chap. 10—Mottier-Lopez and Morales Villabona, and Chap. 11—Smith). Part II contributors would also easily agree with the central role of collaborative learning in the teachers' PD process. Collaborative learning activities act as a trigger of self-awareness and personal reflection and are an important source of feedback on professional learning achievements.

Collaborative activities not only help professional learning to occur but are also crucial in making it durable. They generate resources and practices to improve learning (Chap. 12—Swaffield et al.). Collaborative activities also contribute to a snowball effect: 'As the pool of teachers who are knowledgeable and comfortable with AfL grows, so too do the opportunities to share, learn, and reflect together' (Chap. 9—DeLuca et al.). Collaboration stimulates professional learning through peer modelling first and then through the individual leadership of teachers who have become knowledgeable in AfL and are willing to share their positive experiences.

Collaboration is required to maintain internal coherence among teachers and to ‘precipitate a collaborative culture’ (Chap. 9—DeLuca et al.).

A positive attitude towards colleagues, nonjudgmental approaches, and openness to feedback from colleagues are all necessary conditions for collaborative learning to occur and be maintained (Chap. 9—DeLuca et al.). A form of ‘assessment culture’ (Chap. 16, Part III—Birenbaum) appears to be both a prerequisite and a product of professional collaborative learning. This statement raises the question regarding how likely PD is to succeed in a school environment where such pre-conditions are not present.

Collaborative professional learning among teachers is inclusive and may extend to researchers, external experts, regional education authorities, and the school leadership to varying degrees. This collaborative learning is illustrated in one manner or another in all chapters of Part II. Collaboration among teachers may also spread well outside the school environment and reach out to other schools (Chap. 11—Smith). Although multi-school collaboration presents practical challenges, it also provides great benefits such as in the transition from elementary feeder schools to a specific secondary school: ‘AfL principles transcend differences, provide a common language, and promote coherence’ (Chap. 12—Swaffield et al.).

Notwithstanding the potential of collaborative learning to transcend differences, professional development must also take into consideration individual teachers’ existing practices and address issues that are ‘pragmatically relevant’ (Chap. 10—Mottier-Lopez and Morales Villabona). An AfL policy should be based on a rigorous analysis of where teachers are in terms of PD and what the next step in PD should be (Chap. 5, Part I—Griffin et al.). When there are significant differences in terms of readiness, steps must be taken to ensure that there is a shared knowledge base of prerequisites which all teachers may draw on. Centralized learning sessions, for instance, may help facilitate knowledge mobilization at later stages of PD. Direct instruction works better with teachers who have no previous knowledge or experience of AfL, and it should be used to model AfL strategies that teachers are expected to use with students (Chap. 9—DeLuca et al.).

It appears that collaborative learning and PD have greater chances of success if they target AfL skills within the teachers’ zone of proximal development with respect to professional learning. Without basic assessment literacy, learning about AfL may not be within several teachers’ reach or may not be sustainable. For instance, acquiring a competence in AfL needs the mastery of prerequisite skills such as being able to align assessments with students’ learning progressions. Teachers must understand the ‘big ideas’ (Chap. 13—Ruiz-Primo), see behind the curriculum and develop an understanding of the links between lessons inside a unit as well as of the learning progressions essential to achieve a unit’s learning targets. ‘Curriculum mapping’ (Chap. 13—Ruiz-Primo) is a learning activity that provides opportunities for teachers to deepen their understanding of the curriculum and to properly align, at a planning stage, their assessment instruments with the learning expectations. Although curriculum alignment is not sufficient to ensure competence in AfL, efficient AfL strategies can hardly be developed without a basic mastery of curriculum alignment.

Similarly to the use of AfL to support students' learning and help them improve their marks on summative assessments, AfL may be utilised during teacher PD in conjunction with formal evaluation procedures. A principal supervised by a superintendent, and a teacher evaluated by a school principal, can be accompanied through the use and modelling of AfL principles (Chap. 14—Davies et al.). Just as teacher trainers, experts, and researchers need to model AfL in their PD strategies, school leaders must do the same when evaluating teachers for formative purposes and for decisions about tenure or promotion. For instance, supervisors and school leaders may use AfL principles and strategies as a 'leadership tool' to coach teachers when it is time to co-construct evaluation criteria with teachers and provide them with exemplars of expected standards of performance. The same AfL structure, principles, and strategies may be extended to the coaching and supervision of any group involved in the educational system: 'Using these principles of classroom assessment aligns priority, vision, and action across a school system and as a result, leaders' actions are informed and impactful on student, adult, and system learning' (Chap. 14—Davies et al.).

## 8.8 Moving Forward

To move forward, the planning of PD and collaborative learning must consider individual differences among participants. For instance, certain teachers may benefit more from mentoring, one-on-one coaching or occasional team teaching with a peer, whereas others may enjoy large group conversations and exchanges with peers, students, or with school leaders. Consequently, it is crucial to focus on what is meant by collaborative learning and use the proper and efficient forms of collaboration to support participation and engagement in PD. There are several socio-cognitive and interpersonal processes at play in successful collaborative learning initiatives, such as peer modelling, imitation and role playing, learning through observation, cognitive disequilibrium that occurs through social interactions, individual and collective perceptions of self-efficacy, and the capacity to use feedback and to self-monitor one's own practice.

School leadership also plays an important role in making PD successful. To engage in PD, educators and stakeholders at all levels of an education system must perceive the need to improve their assessment skills. School leaders may stimulate such perceptions by challenging existing assessment practices at the same time they inspire educators to set higher standards for themselves and engage them in their own professional learning. Moreover, school leaders also must balance individual development needs with the skills and abilities that must be acquired collectively to meet school improvement targets. Some form of co-regulation is needed to ensure that PD is collectively coherent and individually relevant.

Professional development, like policy implementation, requires space for co-regulations. Such space is needed because developing a capacity for AfL is complex, and adjustments must be made along the professional learning



progression. Acquiring a *true* competence in AfL, one that extends beyond the use of artefacts or quick tricks, requires middle to long-term planning, appropriate forms of support, and collaborative work. Although certain appropriate assessment skills are a must for teachers to develop a competence in AfL, they will not necessarily lead to the expected student learning outcomes if they are not complemented by pedagogical and curriculum alignment skills: ‘a main recommendation ... is for policy makers and leaders of education to invest in multiple small-scale, long-term projects instead of multiple large-scale, short-term activities’ (Chap. 11—Smith).

It is clear from Part II contributions that PD requires much careful planning and attention to the prevailing conditions. Here is a short list of recommendations that can be deduced from Part II contributions to move forward in using PD to develop the capacity to use AfL at the classroom level and at all levels of the school system:

1. Consider existing practices and address issues that are ‘pragmatically relevant’ for teachers.
2. Set high standards while targeting skills that remain within the participants’ zone of proximal development of professional learning.
3. When there are significant differences in terms of readiness, take measures to create a shared knowledge base of prerequisites before moving any further.
4. Plan a variety of instruction and training methods which allow for a rotation between theoretical and practical knowledge, e.g., alternate PD seminars and classroom experiences to allow teachers to develop new conceptions of assessment.
5. Use collaborative work when favourable conditions are encountered such as a positive attitude towards colleagues, nonjudgmental approaches, and openness to feedback from colleagues.
6. Use supervision as an opportunity for PD, both for the teachers and the supervisor. For instance, use AfL principles and strategies as a leadership tool when it is time to co-construct evaluation criteria and provide exemplars of the expected standards of performance.
7. Use modelling by experts as well as peer modelling to illustrate best practices and professional learning targets.

Although the previous recommendations are based on the development of new capacities in AfL, such capacities remain fragile and need a supportive environment to be sustainable and strengthened within each school. Regardless of how successfully policies and PD were in developing these new capacities among teachers, the school and the classroom environment will be the last hurdle in meeting the challenge of implementation:

The key to successful change is the improvement in relationships between all involved and not simply the imposition of top down reform. The new emphasis in educational change is based on creating the conditions to develop the ‘capacity’ of both organizations’ and individuals to learn. The focus moves away from an emphasis on structural change towards changing the culture of classrooms and schools, an emphasis on relationships and values. (Fullan 2002, in Laveault 2008, p. 12)

The chapters of Part III *Assessment Culture and the Co-regulation of Learning* will prolong the reflection already initiated in Parts I and II in the direction described above. The chapters will illustrate how ‘effective implementation of assessment for learning clearly requires the *concerted coordination of policy, professional development, and practice in classrooms and schools*’ (see Chap. 15 for an introduction to Part III).

## References

- Brookhart, S. M., Moss, C. M., & Long, B. A. (2010). Teacher inquiry into formative assessment practices in remedial reading classroom. *Assessment in Education: Principles, Policy and Practice*, 17(1), 41–58.
- Butler, D. L., & Schnellert, L. (2012). Collaborative inquiry in teacher professional development. *Teaching and Teacher Education*, 28, 1206–1220.
- Carter, K., Sabers, D., Cushing, K., Pinnegar, S., & Berliner, D. C. (1987). Processing and using information about students: A study of expert, novice, and postulant teachers. *Teaching and Teacher Education*, 3(2), 147–157.
- Earl, L., & Timperley, H. (2014). Challenging conceptions of assessment. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning*. Dordrecht: Springer.
- Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment. *Educational Measurements. Issues and Practice*, 28(3), 24–31.
- Herman, J. L., Osmundson, E., & Silver, E. (2010). *Capturing quality in formative assessment practice. Measurement challenges*. CRESST Report 770. Los Angeles: CRESST/University of California. <http://files.eric.ed.gov/fulltext/ED512648.pdf>. Accessed October 24, 2015.
- Hill, M. F. (2011). ‘Getting traction’: Enablers and barriers to implementing assessment for learning in secondary schools. *Assessment in Education: Principles, Policy and Practice*, 18(4), 347–364.
- Ingvarson, L. (1998). Professional development as the pursuit of professional standards: The standards-based professional development system. *Teaching and Teacher Education*, 14(1), 127–140.
- Laveault, D. (2008). Mesure sans démesure: la contribution de Jean Cardinet aux méthodologies de la mesure et de l’évaluation en éducation. *Mesure et évaluation en éducation*, 31(2), 5–17.
- Looney, J. W. (2011). *Integrating formative and summative assessment. OECD Education Working Papers, No. 58*. Paris: OECD Publishing.
- Parr, J. M., & Timperley, H. S. (2010). Multiple ‘black boxes’: Inquiry into learning within a professional development project. *Improving Schools*, 13(2), 158–171.
- Reynolds, R. J. (1970). Classroom verbal interaction patterns as a function of cognitive complexity. *The Journal of Teacher Education*, 21(1), 59–64.
- Sato, M., Wei, R. C., & Darling-Hammond, L. (2008). Improving teachers’ assessment practices through professional development: The case of National Board certification. *American Educational Research Journal*, 45(3), 669–700.
- Schneider, M. C., & Meyer, J. P. (2012). Investigating the efficacy of professional development program in formative classroom assessment in Middle School English Language Arts and Mathematics. *Journal of Multidisciplinary Evaluation*, 8(17), 1–24.
- Shepard, L. A. (2006). Classroom Assessment. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 623–646). Westport, CT: American Council on Education/Praeger Publishers.
- Webb, M., & Jones, J. (2009). Exploring tensions in developing assessment for learning. *Assessment in Education: Principles, Policy and Practice*, 16(2), 165–184.

# Chapter 9

## Implementing Assessment for Learning in Canada: The Challenge of Teacher Professional Development

Christopher DeLuca, Adelina Valiquette and Don A. Klinger

**Abstract** In recent years, educational assessment policies in Canada have expanded to include explicit mandates towards assessment for learning (AfL). These mandates emphasize the continuous use of assessments to support student learning through an integrated approach to teaching and assessment. In this chapter, we explore the emergence and implementation of AfL practices and policies in the education systems of several English-language provinces, with a specific focus on the Ontario context. We focus our argument on the critical role of in-service teacher learning for successful AfL policy implementation. Based on a three-year professional learning project in Ontario, we describe and analyze structures to support teachers' integration of AfL within their classroom contexts. Our findings highlight the value and role of responsive, scaffolded, and embedded structures that move teachers toward greater AfL integration. The chapter concludes with a discussion of implications for future professional learning and for policies supporting greater AfL integration in classrooms and schools.

### 9.1 Towards AfL Integration in Canadian Schools

An enduring challenge for education systems is to enact policies, programs, and practices that will optimize the likelihood of student academic success. In 1998, Black and Wiliam captured the interest of educators when they summarized evidence that (a) systematic formative assessment enhanced student achievement, (b) the practice of formative assessment in classrooms could be significantly improved through professional learning, and (c) trustworthy examples of how to

---

C. DeLuca (✉) · A. Valiquette · D.A. Klinger  
Queen's University, Kingston, Canada  
e-mail: cdeluca@queensu.ca

implement and improve the practice of formative assessment were already available. At the same time, they identified a significant problematic in using assessment to promote learning: ‘We acknowledge widespread evidence that fundamental change in education can be achieved only slowly—through programs of professional development that build on existing good practice’ (p. 2).

In response to Black and Wiliam’s (1998) seminal summary of research, other scholars have further explored the benefits of formative assessment on teaching and learning (e.g., Davies 2007; Harlen 2006; Volante 2006). Formative assessment has evolved into principles characterized by the expression Assessment for Learning (AfL). AfL can be regarded as ‘the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go, and how best to get there’ (Assessment Reform Group 2002, p. 2). Given the mounting evidence that formative assessment supports student learning, educational systems have begun to emphasize the role of assessment in relation to teaching and learning within assessment policies.

The evolution of AfL policy in Canada began in the 1980s with a general movement towards using formative assessments to better guide the teaching of students to meet educational expectations across diverse learning needs. In this setting, formative assessment became an essential tool for informing teachers’ instructional practices and for promoting differentiated teaching and learning. At the same time, educational jurisdictions were subjected to increasing calls for accountability. This required teachers to explicitly demonstrate that their efforts were helping to ensure that students were achieving important educational outcomes (Klinger et al. 2008; Ryan 2002; Volante 2007). The foundations for these accountability efforts were commonly linked to educational outcomes, as measured through large-scale assessments. Given the emphasis on assessment, a similar focus began to evolve at the classroom level, resulting in an increased emphasis on formative assessment strategies. The increased attention on classroom and large-scale assessments highlighted the need for teachers to develop their assessment literacy.

Assessment literacy continues to be a core competency for teachers across the nation (Council of Ontario Directors of Education 2006; Stiggins 2002; Volante and Earl 2013), and the value of instituting policies, developing resources, and offering professional development opportunities that support AfL practices continues to grow. As identified in both professional standards and current research, teachers are required to be assessment literate professionals: educators with the knowledge of how to construct, administer, and score reliable assessments and communicate valid interpretations about student learning and achievement (Popham 2004; Stiggins 2002). Assessment literacy ultimately involves integrating assessment practices, theories, and philosophies throughout the processes of teaching and learning. At its most progressive stage, educators within Canadian schools are now creating an assessment culture where the student and teacher learn together in a collaborative relationship predicated on AfL. Nevertheless, it is not clear that these progressive practices are becoming the norm in the majority of classrooms and schools or that in-service teachers are given the professional learning opportunities

they need to develop assessment expertise and implement AfL within their classrooms.

Systemic implementation of new educational practices is a challenge in Canada because education falls under provincial jurisdiction. Each of Canada's 10 provinces and 3 territories is responsible for implementing educational policies and curricula. Hence, the development of AfL-centered policies has not been uniform across the country. Several English-speaking provinces have recently developed policies that aim to promote AfL as a fundamental component of teaching and learning (e.g., Alberta Assessment Consortium 2005; British Columbia Ministry of Education 2004; Ontario Ministry of Education 2010). For example, British Columbia's Accountability Framework promotes 'evidence-based, data-driven decision-making with a focus on assessment for learning' (British Columbia Ministry of Education 2002, p. 1). The British Columbia policy parallels others that articulate an appreciation of how AfL adoption can increase student engagement and support students' growth towards valued educational outcomes. Another example is evident in the Western and Northern Canadian Protocol for Collaboration in Education, which represents Ministries of Education in Western and Northern Canada. This group of provinces and territories published a document in 2006 entitled *Rethinking Classroom Assessment with Purpose in Mind*. The document describes assessment as follows:

Assessment *for* learning, assessment *as* learning, and assessment *of* learning all serve valuable, and different, purposes. It is not always easy, however, getting the balance right. If we want to enhance learning for all students, the role of assessment *for* learning and assessment *as* learning takes on a much higher profile than assessment *of* learning. (Manitoba Education 2006, p. 14)

Across provincial assessment policies, there is an emerging emphasis on the value and benefits of integrating assessment for and as learning into classroom teaching and learning (see [www.caffn.ca](http://www.caffn.ca) for complete listing of policies). These policies emphasize the continuous use of assessment to support not only teachers' instructional practices but also students' learning, through an integrated approach to assessment, teaching, and learning.

Given the diverse geographic spread and the degree of provincial autonomy of educational systems across Canada, in this chapter we focus specifically on the AfL policies and practices within the province of Ontario. In Ontario, there has been a substantial effort to support AfL implementation and adoption. In 2010, the Ontario Ministry of Education issued its provincial assessment policy entitled, *Growing Success: Assessment, Evaluation, and Reporting in Ontario Schools*. The Growing Success document explicitly details the significance of AfL by encouraging teachers to 'provide students with descriptive feedback and coaching for improvement' (Ontario Ministry of Education 2010, p. 28). This policy document further articulates the following essential components of AfL that teachers are required to integrate throughout their teaching practice:

- plan assessment concurrently and integrate it seamlessly with instruction;
- share learning goals and success criteria with students at the outset of learning to ensure that students and teachers have a common and shared understanding of these goals and criteria as learning progresses;
- gather information about student learning before, during, and at or near the end of a period of instruction, using a variety of assessment strategies and tools;
- use assessment to inform instruction, guide next steps, and help students monitor their progress towards achieving their learning goals;
- analyze and interpret evidence of learning;
- give and receive specific and timely descriptive feedback about student learning;
- help students to develop skills of peer and self-assessment. (Ontario Ministry of Education 2010, pp. 28–29)

However, despite the presence of assessment policies focused on AfL, teachers still require significant professional learning support to shift their pedagogical practice towards an AfL orientation. Ongoing efforts to provide such meaningful professional learning highlight the critical need for classroom-embedded, in-service teacher learning opportunities to mobilize new educational policies into effective instructional practices (e.g., Borko et al. 2008; Darling-Hammond and Richardson 2009).

In recognition of the need for more effective models of in-service learning opportunities to support teachers' integration of AfL practices, the Ontario Ministry of Education in partnership with two school boards engaged in a two-year professional learning project aimed at Building Capacity in Assessment for Learning (BCAfL). This BCAfL project represented a partnership with the purpose to support shared and complementary goals: (a) purposeful professional learning about AfL, and (b) enhanced capacity to effectively implement AfL across school board classrooms. The conception of AfL adopted in the BCAfL professional learning project was predicated on the work of Black and Wiliam (1998), which was later formalized as AfL by the Assessment Reform Group (2002) and other scholars (e.g., Stiggins 2002; Black and Wiliam 2006; Wiliam et al. 2004). Representatives from the Ontario Ministry of Education worked collaboratively with the two school boards and external researchers to engage in research that would document the professional learning gains affiliated with this project. The specific purpose of this research was to examine how professional learning structures within the project supported teachers' learning towards greater implementation of AfL policies within their classroom practices. Our findings highlight components of the BCAfL professional learning project that positively supported teachers' learning and implementation of AfL. Overall, these findings support the need for responsive, scaffolded, and embedded structures for teacher learning that explicitly target key AfL strategies within a collaborative framework of learning. Our chapter concludes with specific recommendations for future research and practice aimed at greater AfL integration.

## 9.2 The Professional Learning Project in Ontario

The primary goal of the professional learning project was to support teachers' understandings and integration of AfL within their classrooms. Ultimately, by increasing teachers' capacities to integrate AfL, it was anticipated that students would subsequently develop the knowledge and skills to become more independent, self-monitoring learners leading to increased student achievement. In order to fulfill this primary goal, the two-year professional learning project engaged teachers in (a) eight large-group, centralized learning sessions focused directly on AfL concepts and practices; and (b) sixteen classroom-based instructional rounds to observe and explore the implementation of AfL (City et al. 2009; Marzano 2011). Both of these activities were feedback rich, providing opportunities for teachers to share their evolving understandings of AfL, set learning goals, co-plan AfL implementation strategies, and engage in assessment and monitoring tasks to track their own learning.

Specifically, the instructional rounds component involved observing peer teachers integrate AfL strategies within their own classrooms. The observations were not only guided by a focus of inquiry but also followed by a collaborative debriefing, which consisted of focused conversation about the observation guided by reflective questions and prompts. In preparation for an instructional rounds visit, teachers selected specific AfL learning goals and supported one another in planning for and implementing these goals to support students' learning and adoption of AfL strategies. Through peer observations and assessment consultant debriefing, teachers collaboratively engaged in learning about AfL within their local teaching contexts. Teachers worked directly with colleagues in their own school context, in other schools, and in the neighboring separate school board to engage in the large-group learning sessions and the instructional rounds visits. Further, the Ontario Ministry of Education and school board assessment consultants facilitated and supported these various professional learning opportunities. These support structures served to help teachers interpret AfL theory and policy within their practice. For example, personnel from the Ontario Ministry of Education guided teachers through brainstorming activities on what AfL could look like in their context of practice based on provincial policies and AfL principles. They further engaged teachers in planning lessons and tasks that integrated AfL throughout instruction and then engaged teachers in peer feedback opportunities on their work.

## 9.3 Collecting Evidence on AfL Professional Learning

As part of our collaborative research study about this AfL-focused professional learning project, evidence regarding teachers' changing conceptions and practices of AfL was collected over the two-year period across all teacher participants through varied data collection methods: (a) observations of professional learning

days, (b) classroom and instructional rounds observations, (c) two teacher surveys, and (d) teacher interviews. Data were collected from 88 teachers who represented various backgrounds and degrees of AfL exposure. Participants ranged from being first year teachers to those having over 20 years of teaching experience. Some teachers had participated in a previous professional learning program focused on AfL integration, but for the majority, AfL was a new concept. Teachers were split between teaching divisions (i.e., Elementary—Grades K-8, and Secondary—Grades 9–12).

Observations occurred during each of the centralized professional learning days. These learning days included direct instruction on AfL strategies and the instructional rounds approach, as well as co-planning periods so that teachers could jointly identify opportunities for the integration of AfL within their lesson plans. Observations also occurred during eight of the instructional rounds days with eight subsequent classroom visits. The focus of observation during these periods was on teachers' intended and enacted implementation of AfL strategies. At the end of each year of the BCAfL professional learning project, all 88 teachers were invited to complete an open-ended survey that asked them to describe their understandings of AfL, professional learning goals, implementation strategies, as well as the strengths and weaknesses of the project. At the end of each year, 15 teachers were interviewed to collect additional data on the effectiveness of the BCAfL professional learning project. These semi-structured interviews were approximately one hour in length, held in teachers' classrooms, audio recorded, and transcribed verbatim. Interview questions focused on BCAfL professional learning project components and teachers' previous experiences that supported and limited their learning and implementation of AfL within the project.

Qualitative data from the surveys, observations, and interviews were collectively analyzed using a standard thematic coding process (Patton 2002). From an initial analysis of data, a code list was generated and then codes were grouped into broader thematic categories. Codes with a high degree of co-occurrence (i.e., two or more codes used for same data) were collapsed into broader categories if they represented similar themes. All data were coded by two researchers, with an inter-rater reliability of 96 %. In cases where data were coded differently, raters discussed code assignment until consensus was reached. Naming of thematic categories was negotiated across two researchers in relation to participants' quotations and previous literature. For the purpose of this chapter, we identified five themes that directly related to policies and structures that support teachers' learning and integration of AfL within their classroom contexts.

## 9.4 Supporting Teacher Learning

Five themes became evident related to structures that supported teacher learning and implementation of AfL. These five themes highlight the value and role of responsive, scaffolded, and embedded structures that move teachers along an AfL



learning continuum. The themes are: (a) a networked model for professional learning, (b) direct instruction in AfL, (c) observing assessment in action, (d) engaging in a reflective feedback loop, and (e) supports that maintain momentum for teacher learning.

### ***9.4.1 A Networked Model for Professional Learning***

Teachers in the BCAfL professional learning project found the highly networked learning model that connected them from within and across schools and school boards beneficial. Throughout the project, teachers described the benefits of the ongoing discussions with peers within and outside their teaching contexts about pedagogical issues. Teachers consistently noted that they generated a variety of AfL ideas through these discussions, which became a source of support and resources to promote their AfL goals.

Within the networked model that characterized the BCAfL professional learning project, teachers emphasized structures that were required to maintain a coherent and generative learning experience about AfL. They sought resources that would help them see what successful practices looked like in the contexts of the subjects they taught, and a variety of exemplars that would address specific, yet common, dilemmas in AfL implementation. Specifically, teachers in this project wanted to discuss different AfL approaches, share ideas for ways of providing feedback to students (oral/written, mini-conferences, etc.), and develop methods to help students keep/record/remember/refer to their personal learning goals. One teacher noted, ‘we all struggle with some of the same issues. It is useful for us to talk about how to approach these challenges in our own classrooms.’ Teachers indicated that working and sharing ideas with colleagues were the most helpful practices in subsequently supporting their students’ learning. The result was that these teachers believed that working with partner schools and/or different school boards would be an important component of any professional learning project focused on AfL implementation. The centralized sessions in this project were also considered highly valuable for teachers to negotiate and discuss their various attempts at implementing AfL. There was a perceived value to hearing the experiences of others.

Within the BCAfL professional learning project, teachers emphasized the importance of administrative support in creating a culture of risk taking and learning in the classroom so that they could share their learning with their peers. ‘I can imagine it would be quite a challenge if you had a principal that really wasn’t on board.’ Another teacher continued, ‘because if you’re working without the support of administration, there’s less of a consensus amongst or a sense of collegiality in the implementation of assessment for learning that takes place.’ As teachers became more confident with AfL practices, they were also able to support and promote AfL at the school level with teachers who were not initially involved with the project. ‘Lead teachers have emerged and they are sharing our work with other teachers. This learning is spreading.’ Accordingly, a networked learning

approach precipitated a collaborative culture within the schools and the school boards, which focused on AfL. One teacher recognized this approach by stating: ‘If individuals are provided meaningful opportunities for collaboration with a focus on assessment for learning, then we will further develop a culture of learning/collaboration/shared understanding within our school board.’

Ultimately, the networked learning approach to professional learning contributed towards increased feelings of confidence and competence among teachers as leaders and as learners. ‘If we share our resources, ideas, and strategies as professional learners then we will become more competent in fostering a culture of self-sufficient and independent learners.’

### ***9.4.2 Direct Instruction in AfL***

Our work continues to highlight the value of instruction for teachers to develop new instructional knowledge and skills. Teachers consider direct instruction about AfL a critical component of their learning. These teachers began the BCAfL professional learning project with varying backgrounds in AfL, and the centralized learning sessions helped create a shared knowledge base across teachers. Direct instruction came largely from external experts linked to the project including school boards’ assessment consultants. These experts provided teachers with foundational knowledge and practices in AfL that they could explore more deeply during the instructional rounds observations. As one teacher pointed out, ‘I liked how assessment for learning strategies were used during our workshops that directly modeled for us how to use it in our own classes.’ Teachers suggested that learning and collaborating with peers was enhanced with the inclusion of external experts who not only discussed but modeled AfL for professional learning purposes. There was less focus on expecting the principal or school board personnel to be the ‘resident expert’ and more appreciation for the power of collaborative learning through centralized learning sessions. External experts added motivational and inspirational aspects to learning, especially in this context of AfL, a complex and deeply situational or contextual (e.g., embedded) design for professional learning.

For teachers who were new to the concept of AfL, the direct instruction was critical in ‘getting us started with understanding what it’s all about.’ However, even for teachers who already had many AfL strategies in their repertoire, further direct instruction helped them bring these methods together in a cohesive manner, with an awareness that these methods were all about AfL. Teachers felt they were becoming more purposeful in the strategies they employed in their classroom practices, and they were more readily able to identify, name, and justify AfL in their classrooms. As teachers’ foundational knowledge of AfL continued to increase, they also acknowledged that they were becoming more self-critical, thoughtful, and flexible in their instructional practices. As one teacher stated, ‘because I know more about assessment for learning, I can better tell when I’m doing it and when I could be doing it.’

The increased knowledge of AfL gained from direct instruction directly impacted teachers' planning. As most centralized learning sessions encouraged teachers to identify learning goals for themselves and their students around AfL, they began to use AfL language and priorities in structuring their teaching and learning plans. They ensured that formative assessment activities were embedded in lessons and that students had opportunities to respond to feedback. Teachers noted their increasing efforts to purposefully plan for and pedagogically align, 'using anchor charts,' 'co-constructing success criteria,' 'letting students write the learning goals,' and 'building in peer- and self-assessment activities.' It was evident across the data that direct instruction in AfL had positive impacts on teachers' foundational knowledge to use and plan for the effective use of AfL in their classrooms.

### ***9.4.3 AfL in Action: The Role of Classroom-Based Observations***

Observing AfL in practice through classroom-based observations (i.e., instructional rounds) was highlighted as a key structure that supported teachers' learning in this project. Teachers valued the opportunity to observe their colleagues' classrooms as a way to reflect on and improve their own teaching practices. One teacher noted:

I think the easiest way to improve is to watch someone else because they're going to do things differently than the way you would do it...and seeing those different tactics and strategies helps to keep you from standing still.

Teachers became more confident in their abilities to implement and integrate AfL practices in their classrooms by watching how other teachers were able to use AfL language and strategies in their instruction. Through classroom observations, teachers were also able to observe instances in which teachers struggled to fully integrate AfL. These challenges led to significant learning as teachers were better able to relate to one another through implementation challenges, jointly brainstorming methods for resolving challenges, and ultimately, moving towards a greater understanding of the nuances in implementing AfL. 'By watching other teachers try AfL, and seeing how there are difficulties, it brings us closer together. We can discuss our challenges and work to support each other.'

These teachers also found the debrief sessions after the classroom observations to be a key component to their professional learning in AfL. 'Instructional rounds were very useful. In particular, the debrief sessions during which we shared and discussed our observations.' Teachers embraced a professional collaborative inquiry stance throughout these sessions. They made a concerted effort to improve their classroom practices and appreciated the guidance and suggestions provided by other teachers. One teacher asserted, 'if you want to become better at assessment for learning, perhaps you need to be able to identify what worked and what didn't.' These debrief sessions provided a chance for 'people to come in and say, 'you know what? This was successful, but you could do this to make it better' and 'an

opportunity to ask those questions because it's risk free.' Key to the success of this process was the presence of a nonjudgmental environment that allowed teachers to feel comfortable giving and receiving feedback. The teachers were then very willing to engage in further conversations about AfL integration. Combined, the observations of practices and the end of day debriefs helped to 'close the gap between what I want the students to learn and what I actually think they're learning.'

While the instructional rounds were certainly valuable as a structure to support the integration of AfL in classrooms, our research suggests these instructional rounds may be less effective as an entry point for learning about AfL. We observed that the instructional rounds process was most beneficial for those teachers and administrators who already had a developing or foundational knowledge of AfL. Certainly, the assessment-related activities were richer in those classes in which the teacher had previous experience with AfL through prior school board initiatives. Similarly, the teachers observing instructional rounds classes seemed to benefit more if they themselves had a solid working knowledge of AfL. Previous exposure and experience with AfL, through direct instruction from experts during professional development opportunities, may help teachers become more comfortable having observers in their classroom and enable them to take greater 'risks' to expand their AfL strategies.

#### ***9.4.4 Reflective Feedback Loop***

Teachers involved in the BCAfL professional learning project indicated that one of the greatest positive effects on their instruction and assessment practices came from personal reflection. They consistently expressed a desire for 'time to reflect' on the impact of AfL on their teaching and students' learning. As one teacher noted when asked to think about the impact of the project, 'It is surprising how everything really connects.' Teachers' involvement in the project enabled them to question 'things that I wouldn't have been questioning before...' and to 'have grown significantly in my professional learning.' The teacher data provided evidence of ongoing teacher growth and reflection about their instructional practices related to AfL, and this growth became a powerful element of change in teachers' professional practices. For example, teachers recognized they were experiencing implementation challenges in different classroom contexts and at different grade levels. 'All students were able to use success criteria to give feedback, however, I found that the junior and intermediate students became more skilled at this.' Another teacher pointed out, 'it's easier in elementary in this respect. In high school you only have four months, and the process requires the students to be here when you're generating the success criteria.'

In addition to reflecting on the teaching and learning occurring in their own classrooms, teachers also expressed a need for time to receive feedback on their reflections from colleagues. Personal reflection became more powerful when it was

shared with others, and feedback was given that responded directly to the teachers' initial reflections. Teachers expressed a need for release time to practice what they were learning and to discuss their personal reflections with their peers. 'Just as students need lots of opportunities to practise their skills, so the teaching staff needs the same opportunities to practice to increase their comfort level.' Reflecting on how the project had an impact at the school level also enabled leaders to make decisions that would continue to spread AfL learning at their school. 'Our plan is to extend this more formally next year and have those teachers train others on staff who are interested in adopting the AfL framework.'

The spread of AfL learning through this reflective feedback loop at the school level further promoted a culture of assessment and increased assessment leadership in schools. And this leadership was observed in others besides the school administrators. Teachers in the BCAfL professional learning project were taking on leadership roles, modeling assessment strategies, and leading conversations to share their learning about AfL. One teacher pointed out that 'at our school, a number of teachers have become "experts" in AfL practices... at staff meetings and informally, their knowledge and experience is shared with other teachers.' As more teachers became involved in the project, and developed a foundational knowledge and comfort with AfL processes and practices, they were able to share their learning and knowledge with others. 'Our whole staff has been involved in several different AfL inquiries that have allowed varying groups to develop their confidence in implementing AfL strategies.' As the pool of teachers who are knowledgeable and comfortable with AfL grows, so to do the opportunities to share, learn, and reflect together.

#### ***9.4.5 Supports for Maintaining Momentum***

Teachers involved in the BCAfL professional learning project continually spoke of its value in providing opportunities to support and motivate their ongoing professional learning. These opportunities helped to maintain momentum, a key element for them in creating classrooms where AfL strategies would be embedded as a natural part of teaching and learning. While critical, this momentum was also hard to develop and maintain given the myriad of responsibilities teachers face. 'Momentum—I believe this has been the hardest thing to maintain. When everyone is on board it is much easier to keep focused.' Teachers described ideas for maintaining momentum moving forward: 'we've talked about some things like assessment for learning lunches, things to keep us fresh and motivated and to keep our momentum going.'

Teachers understood there was still much to learn and that they would require an accessible support network to assist them in implementing the program to achieve their professional learning goals. One potential advantage of having teachers at different places in their learning was the ongoing desire of the teachers to continue to observe and have discussions with teachers about AfL integration. One teacher

pointed out that ‘there are still teachers who are not using AfL in their classroom and when students transfer to the next grade we have to start at square one.’ Another stressed that there are ‘a number of very strong teachers who are involved in this project and who are available for assessment leadership in the school.’ These teachers looked for opportunities to model the use of AfL principles, and to share with other teachers, in order to maintain momentum in teacher learning. While it was evident that ‘pockets of learning cultures’ had developed with respect to AfL, for these ‘pockets’ to spread into a more systemic culture, the program as a whole and the teachers within the program would require ongoing support. There was a resounding sentiment that for AfL to become systemic, teachers, principals, and school board leaders would have to maintain momentum. They would need to not only deepen their learning with existing teachers in the BCAfL professional learning project but also spread the learning across schools and the school board. ‘We need to keep going and share our learning. Continuation is so important; it’ll embed the AfL process so much smoother if it doesn’t seem like it’s just all these separate pieces, if we keep at it.’

The integration of AfL at the student level (e.g., using success criteria, learning goals, peer feedback, self-reflection) also requires continued classroom support. Based on the teachers’ reflections, students made significant gains throughout the BCAfL professional learning project. Through continuous teacher efforts with AfL (e.g., prompting students to apply feedback), one teacher noted that, ‘if they don’t understand something I find that they’re willing to ask questions because they know it’s an opportunity to really develop; it’s not one shot.’ Another teacher said, ‘At the end they were conditioned to it and success criteria were just firing off. The feedback and the self-assessment all wrapped in nicely together.’ Overall, teachers observed students using feedback more effectively and becoming more comfortable using learning goals and success criteria through persistent attention to these AfL strategies. Teachers noted increases in students’ independence and ownership for their learning and changes in the sense of community in the classroom. Students were having learning-related conversations without teacher involvement and highlighting their growing ownership of their own learning. ‘All of a sudden they’re making the transition now into their learning activities that I am not a part of directly.’ It was evident that students were making significant gains, but as with the teachers, a commitment to providing continued support would be required for AfL to become embedded in students’ learning practices.

## **9.5 Implications for Professional Development and Policy**

The search continues for models of effective professional learning (professional development) and for policies that result in changing teachers’ instructional practice. Certainly, there is general agreement that professional learning must be ongoing and embedded into teachers’ everyday practices. It also requires continued effort and exploration. Nevertheless, such professional learning is likely more

challenging when the focus is on aspects of teaching in which teachers are less knowledgeable and confident. Classroom assessment is one such aspect of concern for teachers, especially with the relatively recent introduction of concepts such as Assessment *for*, *as*, and *of* Learning. Teachers are now expected to integrate formative assessment practices, and at the same time, also help their students use this information to guide their own learning. The intentions of the BCAfL professional learning project were to use current models of professional development to help teachers begin to understand the complexities of classroom assessment, and begin to implement effective formative assessment practices in their classrooms. The centralized professional learning days coupled with the instructional rounds provided the opportunities to meet these intentions.

Our research findings from the BCAfL professional learning project support the use of embedded models of professional learning. Further, our work indicates that such models must develop a long-term strategy, especially when such learning requires the implementation of new and unfamiliar instructional and learning practices. In our case, we were attempting to implement new conceptions of formative classroom assessment. Derived from our findings are the following continued areas of development for future professional learning on AfL:

1. Develop a trusting professional learning environment to enable teachers to take the necessary risks to develop and refine their practices.
2. Recognize that teachers may be at different levels of understanding with respect to new practices, and that this will impact their comfort and confidence in the implementation of these practices.
3. Help teachers develop their fluency in terms of knowledge and practices with AfL strategies (e.g., sharing success criteria and learning goals).
4. Continue to focus on developing knowledge, skills, and practices related to feedback, and peer- and self-assessment, mainly by practicing these skills during professional learning sessions.
5. Ensure teachers, and subsequently students, recognize and can articulate the value of these practices to develop self-regulation skills to support their further learning.

Combined, these findings highlight the value of a collaborative professional learning structure that pairs teachers who are novice learners about AfL with those who are developing leadership capacity in AfL. Interestingly, even those teachers with more extensive AfL skills continued to see themselves at the beginning stages of implementing AfL practices, and were only becoming marginally comfortable promoting their knowledge beyond the confines of the project itself. Within the project, the collaborative structure provided a positive professional learning environment. These teachers looked forward to modeling the use of AfL principles and sharing with other teachers in the project. There was evidence throughout the data of ongoing teacher growth and reflection about their instructional practices related to AfL, and this growth became a powerful element for change in their professional practices. It was evident that ‘pockets of learning cultures’ had developed with

respect to AfL. Nevertheless, for these ‘pockets’ to spread into more systemic AfL practices and further learning, the project as a whole, and the teachers within the project, require ongoing resources and supports, including continued opportunities to observe and co-plan with other teachers. Specifically, there is a need to establish policies that both govern how assessment is practiced in classrooms (i.e., AfL integration policies) but also policies that shape assessment education requirements for teachers. Directing policy development towards these two ends will help facilitate greater integration of AfL in schools through focused and directed professional learning activities. Establishing AfL integration and assessment education policies at ministry, school board, and school levels will encourage a systemic approach to AfL implementation with a greater likelihood of AfL adoption. Further, these policies should consider the multiple roles and responsibilities required for a systemic shift towards AfL; this means, implicating school board administrators and consultants, school principals and department leads, and external support networks in service of teacher AfL learning.

More importantly, the sustainability of projects such as the BCAfL professional learning project also require an expansion plan to move the learning to other educators and schools beyond those involved in the initial project. Along with having the resources to support such expansion, there is likely a need for a communication strategy and policy structure to disseminate the developing knowledge and skills about AfL. Success stories from those involved in the professional learning can serve as a powerful incentive to encourage other teachers to begin to explore these new emerging practices in their own classrooms. While the teachers involved in this project may not yet see themselves as instructional leaders, their experiences and developing knowledge are critical to the sustainability and expansion of valuable professional learning initiatives. These are the teachers who must encourage other, often less interested, teachers not involved in the professional learning to explore these developing instructional practices. Leveraging existing teacher learning and building upon the stories of professional development projects, such as BCAFL, has the potential to shape meaningful policies aimed at teacher learning. Policy development at provincial and school board levels should respond to research on teachers’ preferred methods of professional learning and integrate these methods with an AfL orientation to teacher development. As evident through this research, when teachers practice AfL strategies in their own professional learning they generate a greater understanding of AfL principles and implementation strategies. Coupling AfL practices with provisions for professional development will yield assessment education policies with promising potential to shape teacher practice.

While our research focused on efforts to help teachers develop a foundational understanding of current conceptions and practices around AfL, the findings provide a sound foundation for the professional learning of complex teaching and learning practices. The extended collaborative nature of the project supported teachers’ learning about AfL. There was sufficient time to discuss and make changes to the model. As one example, our research resulted in refinements to the structure of the central professional learning days and the instructional rounds days.



These changes resulted in greater synergy and links between the two activities, while also providing further opportunities to explore and discuss AfL practices at deeper levels. Our remaining challenge, and the challenge for other professional learning efforts related to AfL, is the lack of evidence of long-term sustainability and impact on teachers' practices and students' learning. Such challenges can only be met through ongoing monitoring and review as articulated in school board and ministry level policies aimed at AfL implementation. In the case of the BC AfL professional learning project, we have little evidence regarding the impact of the project on students' learning. Such evidence could be obtained through continued observations of the classrooms of teachers involved in the project. Not surprisingly, as researchers, our desire would be to conduct these observations over an extended period along with the collection of other forms of data. Presenting additional empirically supported examples of AfL implementation is one step forward. We also see value in pursuing the development of systemic policies that support AfL integration in classrooms as well as assessment education for teachers, principals, and educational administrators. In particular, articulating a collaborative approach to assessment learning appears important, with clear roles for leaders and teacher learners. With directed professional learning efforts and targeted policy development, AfL integration across Canadian schools is possible.

## References

- Alberta Assessment Consortium. (2005). *A framework for student assessment*. Edmonton, AB: Author.
- Assessment Reform Group. (2002). *Assessment for learning: 10 principles*. [https://assessmentreformgroup.files.wordpress.com/2012/01/10principles\\_english.pdf](https://assessmentreformgroup.files.wordpress.com/2012/01/10principles_english.pdf). Accessed January 6, 2015.
- Black, P., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. London: Kings College.
- Black, P., & Wiliam, D. (2006). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 81–100). London, UK: Sage.
- Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M. (2008). Video as a tool for fostering productive discussions in mathematics professional development. *Teaching and Teacher Education*, 24, 417–436.
- British Columbia Ministry of Education. (2002). *Accountability framework*. Retrieved from <http://www2.gov.bc.ca/gov/topic.page?id=DCC1C4815C65494E98F53969B67DA3B4&title=Accountability%20Framework>
- British Columbia Ministry of Education. (2004). *Classroom assessment and evaluation*. Victoria, BC: Author.
- City, E. A., Elmore, R. F., Fiarman, S. F., & Teitel, L. (2009). *Instructional rounds in education: A network approach to improving teaching and learning*. Cambridge, MA: Harvard Education Press.
- Council of Ontario Directors of Education. (2006). Consistency in classroom assessment. Retrieved from: <http://www.cpco.on.ca/ProfessionalDevelopment/Resources/CCA-Final.pdf>
- Darling-Hammond, L., & Richardson, N. (2009). Research review/teacher learning: What matters. *Educational Leadership*, 66, 46–53.

- Davies, A. (2007). *Making classroom assessment work* (2nd ed.). Courtenay, BC: Connections Publishing.
- Harlen, W. (2006). On the relationship between assessment for formative and summative purposes. In J. Gardner (Ed.), *Assessment and learning* (pp. 87–103). Los Angeles, CA: Sage Publications.
- Klinger, D. A., DeLuca, C., & Miller, T. (2008). The evolving culture of large-scale assessments in Canadian education. *Canadian Journal of Educational Administration and Policy*, 76. Retrieved from <http://www.umanitoba.ca/publications/cjeap/articles/klinger.html>
- Manitoba Education. (2006). *Rethinking classroom assessment with purpose in mind: Assessment for learning, assessment as learning, assessment of learning*. Winnipeg, MA: Minister of Education, Citizenship and Youth.
- Marzano, R. J. (2011). Making the most of instructional rounds. *Educational Leadership*, 68(5), 80–82.
- Ontario Ministry of Education. (2010). *Growing success: Assessment, evaluation, and reporting in Ontario schools*. Toronto, ON: Queen's Printer.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Popham, W. J. (2004). Why assessment illiteracy is professional suicide. *Educational Leadership*, 62(1), 82–83.
- Ryan, K. (2002). Assessment validation in the context of high-stakes assessment. *Educational Measurement: Issues and Practice*, 21(1), 7–15.
- Stiggins, R. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(1), 758–765.
- Volante, L. (2006). Reducing bias in classroom assessment and evaluation. *Orbit*, 36(2), 34–36.
- Volante, L. (2007). Educational quality and accountability in Ontario: Past, present, and future. *Canadian Journal of Educational Administration and Policy*, 58. <http://www.umanitoba.ca/publications/cjeap/pdf%20files/volante.pdf>. Accessed January 6, 2015.
- Volante, L., & Earl, L. (2013). Standards, accountability, and student assessment systems: An international overview. *Education Canada*, 54(3). Retrieved from: <http://www.cea-ace.ca/education-canada/article/standards-accountability-and-student-assessment-systems>
- William, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy, and Practice*, 11(1), 49–65.

# Chapter 10

## Teachers' Professional Development in the Context of Collaborative Research: Toward Practices of Collaborative Assessment for Learning in the Classroom

Lucie Mottier Lopez and Fernando Morales Villabona

**Abstract** This chapter presents a collaborative research project carried out with six Geneva primary school teachers. The focus of the project was on teachers' practices of collaborative assessment for learning in their classrooms. The main features of collaborative research are presented, in particular the process of co-construction between researchers and practitioners of a significant project for both the scientific and the professional communities. Interplay between professional development seminars and teachers' classroom experiences was at the heart of the project. Support for teachers' learning was provided by the articulation of conceptual tools proposed by the researchers with concrete tools and data coming from the teachers' classrooms. The conceptualization of collaborative assessment for learning in classroom included both individual and group self-assessment procedures in the context of student work in small groups. An overview is given of the principal themes emerging during three professional development seminars and the intervening experiences in the classrooms. One particular theme is developed in order to illustrate the exchanges and issues considered by the participants. This theme concerns the focus of collaborative assessment for learning on social and/or academic objectives and the corresponding assessment criteria. It highlights teachers' representations about collaborative assessment and, more broadly, their stance and sense of their responsibility with respect to assessment of student learning. The chapter's conclusion outlines some recommendations for professional development in the context of collaborative research.

---

L. Mottier Lopez (✉) · F. Morales Villabona  
University of Geneva, Geneva, Switzerland  
e-mail: lucie.mottier@unige.ch

© Springer International Publishing Switzerland 2016  
D. Laveault and L. Allal (eds.), *Assessment for Learning: Meeting the Challenge  
of Implementation*, The Enabling Power of Assessment 4,  
DOI 10.1007/978-3-319-39211-0\_10

161

## 10.1 Introduction

The research literature frequently points to collaboration among professionals as a factor that can sustain professional development at both the individual level (e.g., a teacher's professional skills and identity) and the collective level (e.g., a school as a learning community) (Gosselin et al. 2014). Bergold and Thomas (2012) define several fundamental principles which guide participative and collaborative research approaches. They state: 'A "safe space" is needed, in which the participants can be confident that their utterances will not be used against them, and that they will not suffer any disadvantages if they express critical or dissenting opinions' (our translation). It is also essential to involve the target community directly concerned by the research issue; this means stakeholders are considered as 'co-researchers.' Although different degrees of participation are possible, the determining condition for Bergold and Thomas (2012) is to deeply involve participants in the decision-making process during the research project.

In the field of education, Desgagné (1997) proposes two interrelated goals which characterize collaborative research. The first goal is to support teacher professional development through research. This means 'encouraging the teachers to question and refine their practices and to work together on a wide range of shared problems relating to contemporary education' (p. 36 our translation). The second goal is to provide adequate conditions for the production of scientific knowledge about the teaching practices being studied. Several features of collaborative research are highlighted by the literature, in particular: research questions should be significant for both the scientific and the professional communities; knowing is closely linked to concrete actions; research is seen as a collective enterprise involving co-construction of shared meanings by participants; situated teaching practices are collectively analyzed; critical reflection by both practitioners and researchers is expected in a transformative learning perspective (Bourassa et al. 2007; Vinatier et al. 2012).

This epistemological stance of collaborative research seems particularly relevant for investigating classroom assessment practices in order to better understand professional assessment cultures, teachers' values, the conditions of their authentic practices, and the resources supporting their transformation (Mottier Lopez, in press). We adopt a situated perspective on professional development, seen as closely linked to collective practices of collaboration (Lave and Wenger 1991). In this perspective, individual dimensions (skills, values, identity, personal history) and sociocultural aspects of teacher learning and practice are seen as mutually constitutive. Moreover, the collaborative research group can be conceived as a *learning community* composed of distributed expertise between practitioners and researchers.

This chapter presents a collaborative research project that concerns *collaborative assessment for learning* (CAfL) in the classroom. The expression 'collaborative assessment' can refer to practices developed by groups of teachers outside the

classrooms, including social moderation of assessment judgments (Allal and Mottier Lopez 2014). In the classroom, this expression can refer to students working together to co-construct shared appraisals about individual contributions to group work or about the contribution of the group as a whole to the implementation of the task. In our research, both contexts of collaboration are present: (1) between teachers and researchers in the context of collaborative research, (2) between students, in interaction with their teacher, in CAfL activities in the classroom.

In the spirit of collaborative research, the questions about CAfL must address both scientific and practical concerns. The main questions defined by our research group are the following:

- At the scientific level: The current literature makes increasing reference to collaborative assessment. But what justifies this designation in relation to other well-known forms of assessment involving interactions between students and between students and the teacher? What broader conceptual framework can be developed for CAfL practices?
- At the practical level: Small-group work is encouraged by the school system directives and the curriculum material used by the teachers. What sort of assessment for learning can be developed for situations of small-group work? What are the objectives to be targeted in CAfL? How does the time frame of CAfL fit in with teaching and learning processes? How can CAfL be implemented and managed in the classroom?

In a collaborative research approach, researchers do not have a value-neutral stance, nor an external position. An in-depth relationship between researchers and practitioners is needed to co-construct meanings and to sustain an ongoing dialogue between their respective viewpoints (Desgagné 1997). The following sections of this chapter present the research context, the participants, and some major findings regarding teacher professional development in the context of our collaborative research.

## 10.2 Research Context and Participants

Our research was based on alternation between professional development seminars and teachers' classroom experiences conducted over an entire school year. To participate in the project, the teachers agreed to contribute to the design of new assessment practices and the experimentation of these practices in their classes. This meant: (1) during the seminars, participating in the co-construction of a shared framework for developing new classroom practices, (2) hosting a researcher (second author of this chapter) in the classroom to observe the practices experimented, (3) holding discussions with this researcher outside the seminars in the form of research interviews, (4) accepting that the assessment practices observed be collectively analyzed and discussed during the subsequent seminar, (5) and starting the

**Table 10.1** Teacher participants

	School	Grade	Gender	Years of experience
Teacher 1	1	6	Female	1
Teacher 2	1	7–8	Female	35
Teacher 3	1	7–8	Male	24
Teacher 4	1	7–8	Female	20
Teacher 5	1	GNT	Female	1
Teacher 6	2	5–6	Female	10

cycle again. This alternation aimed at creating conditions for in-depth exchanges, including possible socio-cognitive conflicts between participants, as well as negotiation of new meanings linked to experiences carried out in the authentic environment of teaching practice. The seminars were conducted by the two authors of this chapter, but the classroom observations and interviews with the teachers were carried out by the second author in the context of his ongoing doctoral research (Morales Villabona 2013).

The research was conducted in the context of the second cycle of primary education in the canton of Geneva (grades 5–8: 8–12 year-old students<sup>1</sup>). Each year, Geneva primary school teachers have to participate in 14 h of professional development activities which they can choose from a catalogue of offers. These 14 h take place during school hours. We proposed an offer entitled ‘Classroom assessment and group work.’ In the description of our offer, we formulated the following questions: What are the different formative assessment procedures that can be envisaged for student work in small groups? To what extent can these assessment procedures support students’ skill in assessing themselves or their peers when working in groups? How can group work and student learning be assessed? Our proposal also explained both the professional development and research goals of the project, and the conditions of participation mentioned above.

Table 10.1 presents the characteristics of the six teachers who chose to participate in our project. Five of them were classroom teachers while one was responsible for providing pedagogical support to classes in her school (designated as GNT in the table).

The 14 h of the seminars were distributed over the first semester of the school year, with one full day to initiate the project (September 2013), followed by two half-days (October 2013, January 2014). Classroom experiences were observed between the seminars and the observing researcher prepared a support document for the discussions in the second and third seminars. During the second semester of the school year, a long-term observation was conducted in the classes of each of the

---

<sup>1</sup>In the canton of Geneva, the first grade of kindergarten is designated as grade 1. This means that grades 5–8 correspond to grades 3–6 in the K–12 systems in other countries.

five classroom teachers. A half-day review brought the project to a close at the end of the school year (June 2014).

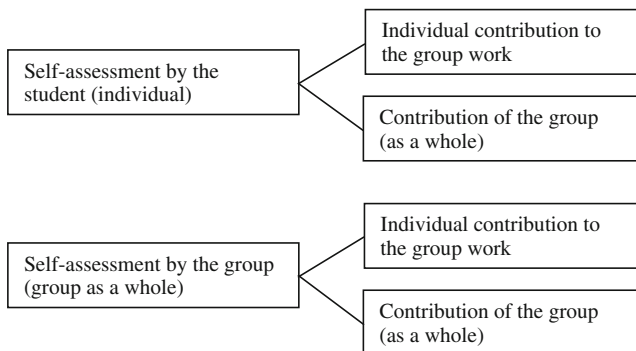
### **10.3 Professional Development Seminars Articulated with Classroom Experiences**

This chapter focuses on the activities carried out in the first semester. We recorded all verbal exchanges and we collected written documents produced during the three seminars. For each seminar, we formulated a 'synopsis' based on the methodological principles defined by Schneuwly et al. (2006). Using this tool, we identified the main themes discussed during the seminars, their succession, duration, and hierarchical structure (themes and subthemes). The notion of theme refers to the objects of concern which emerged from the process of developing a local understanding shared by the participants (Voigt 1985). We transcribed excerpts of significant interactions in the negotiation of collective meanings of CAfL and interpreted these excerpts through an 'analysis by conceptualizing categories,' as defined by Paillé and Mucchielli (2012). Appendix summarizes the results of these analyses for each seminar: (1) the succession of themes of discussion, (2) the decisions taken collectively at the end of the seminar concerning the experiences to be conducted in class, (3) the teachers' own initiatives outside the seminars.

#### ***10.3.1 Rationale of Professional Development Seminars Articulated with Classroom Experiences***

As highlighted above in our epistemological stance, we wished to investigate CAfL in collaboration with the teachers. Our purpose was not to offer a predetermined model but to co-construct shared principles based on both scientific knowledge and teachers' knowledge, in a distributed expertise perspective (Salomon 1993). The principles co-constructed by the members of the research group constituted a negotiated framework that could still be reviewed as the classroom experiences progressed and were collectively analyzed. In this sense, our approach followed the argument by Lussi Borer and Muller (2014) that, in the context of teacher education: 'prescriptions and knowledge should not be transmitted as such, but as objects to be re-normalized, in other words as rules that are resources for action and need to be tested through action and revised (if necessary) according to their viability' (p. 66, our translation).

Each seminar was conducted according to a scenario we had prepared. The first seminar lasted one full day in order to discuss the research orientation and develop



**Fig. 10.1** Self-assessment procedures when performing a task in small groups

the initial professional development questions. It started by an activity of collaborative drafting of a text, in groups of three. The teachers had to co-write a short article that would be suitable for publication in a professional journal<sup>2</sup> on the topic: ‘assessment and student group work.’ The purpose of this activity was to allow the teachers to experience the process of collaboration when carrying out a complex task, including reflections on the implications for collaborative assessment for learning. We proposed several resource documents about the forms of self-assessment (see below, Fig. 10.1) and about co-writing procedures. On the basis of this activity, teachers and researchers began to co-define what CAfL in classroom might concretely represent. Initial shared principles for CAfL were co-constructed and first decisions concerning the experiences to be conducted in class were taken.

All the tested classroom activities were based on the teachers’ existing practices, taking into account what they felt they could achieve, given the official school curriculum and their classroom contexts. This choice was justified for two reasons: (1) to ensure strong ecological validity of the data; (2) to take fully into account the teachers’ professional knowledge. Nevertheless, some constraints were also collectively adopted regarding the common features of the classroom experiences: namely that an academic task chosen by the teacher would be carried out by students in small groups and that a procedure involving a collaborative self-assessment tool would also to be implemented by the students.

In addition to these general constraints, we provided an open-ended planning tool for the design of the classroom activities. It listed the main aspects to be considered by the teacher: the academic activity chosen, the objectives and assessment criteria, the way the students are involved in assessment procedures, the teacher’s role, and other new elements to be experimented. This tool can be considered as an ‘affordance’ (Reed 1996) for the collective discussions and for

<sup>2</sup>The journal (*L’educateur*) is well known by teachers in French-speaking Switzerland.



designing CAfL practices. Its role is both to constraint and to support reflective activities, decisions to be taken, and their regulation.

The teachers, on their own initiative, decided to meet outside the seminars or to communicate by e-mail in order to work together on the development of the assessment procedures to be implemented in their classrooms (see Appendix). We consider these teacher initiatives as a sign that the project was significant for them, that they saw themselves as authors and partners of the common project.

The second and third seminars were conducted in a similar manner: first, a time for the teachers to freely share the assessment practices carried out in their classrooms; then, structured discussions based on a document prepared by the observing researcher. This document included transcribed excerpts of students' interactions during group work and when carrying out CAfL procedures, as well as the different assessment tools developed by the teachers. By presenting these data, our purpose was to allow teachers to acquire a new perspective on their practices and to engage in critical collective thinking. Thus, questions were refined and new shared principles for CAfL were developed, which in turn fed into new classroom experiences, in a continuing process.

### ***10.3.2 Conceptual Orientation of the Seminars and Themes of Collective Discussions***

During the period 1995–2005, a key educational reform was introduced in the canton of Geneva. It played an important role in introducing primary school teachers to the aims of formative assessment, from the involvement of students in the assessment process to the importance of incorporating assessment into daily teaching and learning activities. The practices associated with different forms of self-assessment and with the regulation of learning through formative assessment (Allal and Mottier Lopez 2005) are largely covered in the teachers' initial training, and are then revisited during professional development activities. Although it cannot be assumed that these elements are fully integrated into the practices of Geneva primary school teachers, it is possible to consider that they are part of their professional culture. In this context, we deliberately chose to orient the seminar discussions towards formative assessment procedures involving student peer groups. This choice was coherent, firstly, with the curriculum material used by the teachers, which emphasizes small-group work by students, and secondly, with the primary school directives which do not authorize summative assessment of student group work.

In our project, collaboration between students in the classroom context concerned (1) the academic task carried out in small groups, and (2) the self-assessment process undertaken by the students. The conceptual tool we elaborated for the teachers was based on the idea that collaborative assessment 'can include self-assessment by individuals or by the group as a whole of the product they have

generated, and/or their respective contributions towards the product' (Race 2001, p. 5). Figure 10.1 presents the tool we proposed to the teachers. Two main levels are differentiated:

- The first one refers to self-assessment by a student concerning the individual or collective contributions to the task carried out by the group;
- The second one refers to self-assessment by the student peer group concerning the individual or collective contributions to group work. In this case, self-assessment (by the group as a whole) requires dialogue, the confrontation of different viewpoints, and the construction of shared appraisals.

Although the teachers had some experience with the first level, the second level was new for them.

This conceptual tool was aimed at supporting the design and the experimentation of different CAfL procedures in class. It was designed to be simple and straightforward so that the teachers could easily appropriate the categories. It allowed them to share a common language, to be able to designate the self-assessment level (individual versus group) under consideration, and to collectively imagine concrete examples and practical modalities of CAfL.

Appendix shows the progression of the discussion themes and the detailed questions emerging through collective reflections in the seminars, in relation with the classroom experiences and the data gathered by the observing researcher. Table 10.2 summarizes the principal CAfL dimensions that were particularly relevant for the research group: (1) the individual and group levels of CAfL, (2) the criteria defined in CAfL tools and their uses, (3) the social organization of CAfL (small-group work and whole-class discussions). Table 10.2 also mentions decisions taken regarding successive classroom experiences (in the table: *For Cl-Exp*) and the new questions (*NQ*) resulting from the dynamic interplay between collective discussions and classroom experiences.

Starting with the initial shared idea that CAfL should have a formative function, the research group was essentially concerned with the kind of student learning that CAfL should support (academic and/or social skills) and with the challenge of designing collaborative self-assessment procedures for students working in small groups. Technical and procedural aspects then had to be considered: What sort of tools can be constructed? Which criteria could best support student collaboration? How to use criteria with the students? Substantive issues linked, for instance, to the regulation of student learning became significant only after the first classroom experiences and collective discussions.

### 10.3.3 The Type of Learning Assessed: Social and/or Academic Skills?

To conclude this section of the chapter, and in order to illustrate how the exchanges unfolded, we will discuss one particular theme which was recurrent throughout the seminars. This theme concerns the choice of the objectives targeted by CAfL. Across the three seminars, an evolution was observed in the teachers' stance and sense of their responsibility toward assessment of student learning.

**Table 10.2** Principal CAfL dimensions discussed and experienced in the classroom

CAfL	Seminar 1	Seminar 2	Seminar 3
Individual/Group levels of CAfL	Formative assessment	Focus on group level only	<i>NQ</i>
	Collaborative assessment for learning	Reflections on conditions for CAfL	Does CAfL contribute to regulation of learning? What time frame should be adopted for CAfL?
	Encourage the students to construct shared appraisals		
<b>Criteria in CAfL tools and use of tools</b>			
Academic/Social skills	<i>For Cl-Exp</i> Focus of CAfL on students' social skills and group functioning	<i>NQ</i> Which criteria should be adopted to support student collaboration?	<i>NQ</i> Which learning objectives should be the focus of CAfL? How should academic and social objectives be interwoven?
			<i>For Cl-Exp</i> Including academic objectives in addition to group functioning Remaining focused at the group level
Use of tools	<i>For Cl-Exp</i> Criteria at both the individual and the group levels Using the tool immediately after the academic activity carried out in small groups	<i>NQ</i> Why and how should students play a greater role in defining the criteria and constructing the tools?	<i>NQ</i> What is the role of the assessment tool? At which point in the activity should it be used? What are the benefits of collective construction (by students and teacher) of the assessment criteria?
		<i>For Cl-Exp</i> Criteria to be defined interactively with the students Appraisal will be communicated by open-ended comments written by the group	

(continued)

**Table 10.2** (continued)

CAfL	Seminar 1	Seminar 2	Seminar 3
<b>Social organisation of CAfL</b>			
Small-group work	<i>For Cl-Exp</i> Activities to be conducted collaboratively by small groups of students	<i>NQ</i> How do the students collaborate in CAfL? What do students refer to when constructing group agreement during the assessment procedure? What are the modes of resolution of disagreements?	<i>NQ</i> What is the scope for individual reflection during CAfL?
		<i>For Cl-Exp</i> Work with the students on handling possible disagreements within the group	
Whole-class discussions	<i>During Cl-Exp</i> Some teachers initiated a whole-class discussion about the assessments carried out	<i>NQ</i> What is the role of whole-class discussions with regard to CAfL?	<i>NQ</i> What are the contributions of small-group moments and whole-class moments with regard to CAfL?
		<i>For Cl-Exp</i> Carrying out whole-class discussions before and after CAfL procedures	<i>For Cl-Exp</i> Continue to carry out whole-class discussions

### 10.3.3.1 First Seminar

At the first seminar, during the collaborative writing task, the teachers asked themselves how to articulate transversal<sup>3</sup> objectives (in particular, social skills) and academic objectives in order to assess group work: How can the development of social skills be combined with the acquisition of academic knowledge? Should one be favored over the other in the context of CAfL? The teachers' opinion was to begin by supporting the development of social learning in group work in order to create adequate conditions for academic learning. One teacher stated that 'it perhaps makes sense to teach the students to work in groups first before subsequently introducing learning [academic objectives]' (teacher 3, school 1).

<sup>3</sup>In the curriculum of French-speaking Switzerland, the term 'transversal' objectives refers to objectives that are pursued in all disciplines, as contrasted with academic objectives that are specific to a given discipline.

As researchers, we participated in this debate by stressing that this issue was very relevant, including from our scientific perspective. While not saying whether it would be preferable to begin by one or the other, we highlighted that transversal and academic objectives are closely interrelated in the situated learning perspective we adopt. We explained that, in this perspective, the conditions in which knowledge develops (here, the social forms of student participation in small-group activities) are seen as an integral part of what is learned (Brown et al. 1989). Consequently, both kinds of objectives should be included in CAfL concerning an academic task carried out in small groups. The teachers seemed not to be totally convinced by our researcher viewpoint and theoretical argumentation. For them, if the aim for the students is to collaborate in accomplishing an academic task, CAfL must first target the social skills required. Thus, one teacher stated that ‘self-assessment will concern the collaboration between the students more than the content [academic learning] ... we can include elements of the content but they don’t have to do everything’ (teacher 1, school 1). The teachers’ worry was to avoid overloading the students with the two kinds of objectives. At the end of this first seminar, the research group chose to focus on self-assessment procedures at both individual and group levels concerning how students work together when performing an academic task.

### 10.3.3.2 Second Seminar

At the beginning of the second seminar, the teachers were invited to express their impressions about their first classroom experiences, with which they were relatively satisfied. They were especially pleased with the degree of autonomy shown by their students during the group work and the assessment procedures. Assessment tools constructed outside the first seminar were presented. Most of the assessment criteria defined by the teachers focused on ‘participation,’ ‘group functioning,’ and ‘group collaboration.’ The way the students indicated their appraisal with respect to each criterion varied between the classes: a four-point frequency scale was proposed in some classes, a dichotomous ‘yes/no’ scale, with spaces for open-ended commentary, was proposed in other classes.

Some criteria called for an individual self-assessment about one’s own participation or cognitive contribution, for example:

- I listened to the ideas of my classmates.
- I proposed sentences for the text we are writing.

Other criteria solicited self-assessment at the group level, essentially about contributions of the group as a whole, for instance:

- We listened to everybody’s opinion.
- We spoke softly so we did not disturb the others.
- We avoided off-task talk.
- We worked effectively as a group.

A few criteria were about cognitive aspects required by the academic task, for example:

- I feel able to explain the two themes we worked on.
- Each member of the group practiced explaining the content of the reading.

We noted that the different levels distinguished by the conceptual tool (Fig. 10.1) were present in the experimented classroom practices and seemed to be relevant.

During the seminar, the teachers were quite critical of two principal aspects of the classroom experiences. First, they regretted that the assessment procedures they had tried out did not allow them to gather any information about ways in which the groups sometimes did not function well; they felt that such information could be useful for formative interventions aimed at regulating learning progress. We took advantage of this observation to initiate a discussion about the role of disagreements between students while carrying out both the task and the group self-assessment: By which means would it be possible for the students to resolve these disagreements? Is it always necessary to come to an agreement? The research group finally considered that this aspect deserved further exploration through new classroom experiences, in particular to obtain a better understanding of the potential interactive regulations between students in CAfL. The second aspect concerned the teachers' disappointment about most of the group work products (the students' texts). Although the teachers were interested in the interactional processes between students in CAfL, they nevertheless kept an eye on the academic product about which they formulated their own judgment.

After the exchanges about the first classroom experiences, the group examined some excerpts of peer interactions during group work, prepared by the observing researcher. These excerpts allowed the teachers to discover part of the content of the exchanges between students and to reflect on the potential value of student interactions in CAfL. One teacher stated:

On reading the excerpts, we discover the wealth of interactions between students, which is not easy for us to see directly on the [assessment] tool. We clearly see that the students' reflections on their own work in the group are included (teacher 3, school 1).

The excerpts of student interactions were especially appreciated by the teachers because they provided access to information which they would not have otherwise been aware of.

During the classroom experiences, some teachers initiated whole-class discussions about the assessments the students had carried out. The research group discussed the importance of whole-class interactions in order to construct shared meaning with the students about the new assessment practices and criteria. Progressively, whole-class discussions were seen as an integral part of the design of CAfL. As one teacher stated, 'for the young students, the whole-class discussions can be more valuable than to fill in a chart' (teacher 4, school 1). And in another teacher's opinion, 'the tools are too abstract for the students ... because it is difficult for them to reflect on what they did, to put that into words, and to argue' (teacher 1,

school 1). Based on the first classroom experiences, it was decided that the assessment criteria to be included in the tools must be co-defined with the students during whole-class discussions in order to be more significant. Consequently, the teachers adopted a new format for the self-assessment tools. The assessment criteria defined in the whole-class discussions would be copied by each group on the assessment format. After completing the task, the students would write an appraisal of the group's work with respect to each criterion. They would also answer a question about possible disagreements during group work and the assessment process. The teachers' intention was still to focus on students' collaboration skills, seen as being at the heart of CAfL.

### 10.3.3.3 Third Seminar

Three months have passed since the last seminar, so it was more difficult for the teachers to describe their CAfL classroom experiences. The document prepared by the observing researcher allowed them to rediscover the activities they had proposed to the students. The teachers talked about their experiences of co-definition of assessment criteria during whole-class discussions held before students worked on the academic task. Three or four assessment criteria included in each tool were collectively decided with the students in the different classrooms. Examples were:

- Everyone expressed his or her opinion.
- We discussed calmly to make our decisions.
- There was not only one single leader.
- Everybody participated actively in the task.
- We spoke softly and kindly.
- We talked mostly about the task.

In general, the criteria focused only on the social dimensions of the students' activity, but a few criteria mentioned the academic task, for instance:

- We took into account everybody's ideas to write our text.
- We discussed about our writing.

Each group of students wrote a shared appraisal for each assessment criterion. After the use of the assessment tool, whole-class discussions were again held to analyze the assessments carried out by the peer groups. In some classes, several groups of students did not function well. Since the peer groups were not always capable of dealing with this on their own, the concerned teachers decided to talk about these problems during the collective discussions held after the activity. Whole-class discussions thus represented a means of potential regulation of the quality of student interactions.

More generally, the teachers started to express more critical reflection about the idea of CAfL: What is the purpose of assessing group interactions in relation to the academic objectives of the task? How should these different levels be interwoven?

What are the benefits of collective construction (by students and teacher) of the assessment criteria?

At this point, we called attention once again to the fact that assessment criteria regarding the academic objectives were lacking in the assessment tools. A discussion emerged about the possible lack of coherence between the academic task the students had to carry out (focused on writing) and the criteria included in the tools (focused on collaboration skills without an explicit link to writing). The teachers expressed uncertainty about their choices. As one teacher stated:

We invented things so different from our day-to-day experiences in the classroom. Normally, our main aim in assessment is the product of group work. The assessment of group functioning is generally of secondary importance. (teacher 2, school 1)

After two cycles of classroom experiences and critical discussions about them, the teachers appeared to be ready to adjust their representations. It seemed that they first needed to try out CAfL procedures focusing on social skills before deciding if they could also be used to assess academic learning. The teachers found it relatively easy to envisage CAfL targeting transversal objectives (which do not lead to grades), but they were unsure of its usefulness concerning academic learning. As one teacher stated:

I am not convinced by collaborative assessment with regard to academic learning.... concerning transversal aspects yes, but I am yet to form an opinion with respect to academic aspects. (teacher 3, school 1)

To a certain extent, the teachers found it difficult to entrust the assessment of academic learning to the peer groups and more generally to the students. As a teacher stated, ‘we asked the students to do something that I think is part of the teacher’s job ... to play a role which is not their own, so that is not easy’ (teacher 4, school 1). Nevertheless, at the end of the third seminar, the teachers agreed on the need to introduce academic objectives in CAfL procedures, in relation with substantive issues: What is the purpose of collaborative assessment? Does it contribute to the regulation of learning? We noted that these essential questions became significant only after the two cycles of classroom experiences and collective critical reflections. New classroom CAfL practices were planned precisely to explore these questions. The ongoing doctoral research by Morales Villabona will provide results about the outcomes.

## 10.4 Discussion

We think that even the most attractive assessment model will be doomed to failure if it cannot adjust to the constraints and practices of the field. Reports on educational innovation show how difficult it is to implement assessment reforms in teachers’ classrooms (e.g., Gilliéron Giroud and Ntamakiliro 2010). Participative



and collaborative research approaches seek to forge closer ties between the scientific and professional communities. The goal is:

To create an intersection between the two working cultures in order to build a common culture, derived from this process of mediation, where knowledge is constructed in collaboration and takes into account both the constraints and the resources of the two worlds, that of research and that of practice. (Desgagné 1997, p. 383, our translation)

The challenge for researchers is to be able to create conditions for integrating teachers' viewpoints (and the contexts in which teachers practice) with their own scientific frameworks of investigation.

Our chapter has shown how a common project was initiated and developed, regarding CAfL practice, which was a new concept for the participating teachers. It was important to identify and address issues that were pragmatically relevant to the teachers in the context of their assessment practices. Starting with these issues, a deeper understanding was gradually co-constructed between teachers and researchers. Tools, as artifacts, played an important role of mediation between the scientific and the professional communities, whether conceptual tools proposed by the researchers or practical tools and data coming from the classrooms. More significantly, the interpretative activity fostered by these tools, in the setting of professional development seminars and classrooms experiences, led to negotiation of collective meanings and potential transformations of practices.

Our project approached professional development seminars alternating with classroom experiences from a situated perspective. In this view, learning is conceptualized as a transformation of the processes of participation in socially organized activities (Lave and Wenger 1991). The research group, as a community of learning, offered structured collaborative activities favorable to teachers' professional development. The collaborative research project asked the teachers to be 'boundary crossers' (Engeström et al. 1995) able to explore new classroom assessment practices both in deep discussions with researchers and in interaction with the students in their classrooms. The challenge for the teachers was to create new practices based on both experiential and conceptual knowledge. For the researchers, the challenge was to strike the right balance between the scientific and the practical worlds and to maintain favorable conditions for co-regulation between them.

From our experience conducting collaborative research projects (this chapter, Mottier Lopez et al. 2010, 2012; Mottier Lopez 2015), several implications for teachers' professional development can be drawn. In these projects, participants need to develop a shared culture of collaborative inquiry related to their professional concerns. The co-construction of shared values, norms, and collaborative practices takes time. This process seems to go faster when teachers are from the same school, particularly if they are used to working together. In this case, the teachers share a common school culture and specific issues linked to its context. They also have more opportunities to meet outside the formal seminars to pursue professional development projects. There are, however, some advantages of working with teachers coming from several different schools due to opportunities

for confronting different practices and school assessment cultures. Exchanges may be richer, leading to expanded collective and critical questioning between participants. But more time is needed in order to build a relationship of trust and to construct shared meanings within the collaborative research group.

The principal limitation of the project presented in this chapter was its rather short duration (three professional development seminars totaling 14 h, plus the intervening classroom experiences) in order to develop effective collaboration between researchers and teachers and initiate new assessment for learning practices. We think that the duration of collaborative research projects is an important factor in fostering the development of new classroom assessment practices. Several successive cycles of alternation between seminars and classroom experiences need to be implemented, as was the case in a project where collaborative research was conducted over a three-year period (Mottier Lopez et al. 2010). Given the substantial involvement of teachers in collaborative research, it is important that the professional development seminars be carried out during school hours with the support of the school administration which provides funding for release time. A crucial condition for the success of this kind of project is that the school authorities adhere to this form of professional development linked to participation in research.

It would be misleading, however, to idealize collaborative research. It appears that some teachers are at ease with individual forms rather than collective, school-based forms of professional development (Gosselin et al. 2014). In terms of educational policy, we think that it is important to design collective projects of professional development that are articulated with courses to which teachers can sign up individually. In a lifelong learning perspective, we believe that it is crucial for school systems to propose various perspectives and activities for supporting teachers' professional development in assessment.

### **Appendix: Organization and Orientation of the Seminars and the Collaborative Assessment Experiences in the Teachers' Classrooms**

Seminar 1 (7 h)	Seminar 2 (3 h)	Seminar 3 (4 h)
Discussion of the "contract" between teachers and researchers: professional development goals and research goals	Teachers share their practices, carried out in class and observed by the researcher	Idem
Teachers' questions regarding the theme of the seminars: "assessment and student group work", an assessment for learning	Based on a document prepared by the observing researcher	Idem

(continued)

(continued)

Seminar 1 (7 h)	Seminar 2 (3 h)	Seminar 3 (4 h)
Initial activity Collaborative drafting, in groups of 3, of an article for a professional journal on the theme of the seminars; teachers experience directly the process of collaboration when carrying out a complex task	Discussion of excerpts of interactions between students during group work and when using assessment tools: How do the students collaborate? What conditions appear necessary for collaborative assessment?	Discussion (idem): What is the purpose of assessing group functioning in relation to the academic objectives? How should these different levels be interwoven? What are the contributions of group moments and whole-class moments with regard to collaborative assessment and student learning?
Discussion of this experience, with reference to concrete examples in relation to the conceptual framework proposed by the researchers: self-assessment (SA) procedures in student group work (Fig. 10.1)	Discussion concerning the assessment tools created by the teachers and their use in class: What is their role? What conditions will allow these tools to be genuinely conducive to collaborative assessment by students?	Discussion (idem): What are the benefits of collective construction (by students and teacher) of the assessment criteria?
Collective reflection on – assessment criteria (academic objectives/group functioning) – SA at an individual level and at a group level – roles of the classroom teacher depending on his/her intentions Co-construction of shared principles for collaborative assessment procedures in the classroom – ask the students to focus their assessment on social skills and group functioning – encourage the students to construct a shared appraisal during their joint assessments	Refining the questions raised by the participants – What are the role and contributions of whole-class discussions with regard to collaborative assessment and student learning? – What do students refer to when constructing group agreement during the assessment procedure? What are the sources and the modes of resolution of disagreements? – Which criteria should be adopted to support student collaboration? Why and how should students play a greater role in defining the criteria and constructing the tools?	Refining the questions (idem) – Which learning objectives should be the focus of collaborative assessment? Can different types of learning be assessed with the same procedure or tool? – What is the role of the assessment tool? At which point in the activity should it be used? – What is the scope for individual reflection during collaborative assessment? – What time frame should be adopted for collaborative assessment (occasional, continuous, etc.)? – What is the purpose of the collaborative assessment? Does it contribute to regulation of learning?
<b>Decisions concerning the experiences to be conducted in class</b> Three different academic activities are planned	<b>Decisions concerning the experiences to be conducted in class</b> A single academic activity is planned in all classes (text production)	<b>Decisions concerning the experiences to be conducted in class</b> Idem (text production)

(continued)

(continued)

Seminar 1 (7 h)	Seminar 2 (3 h)	Seminar 3 (4 h)
Activities to be conducted collaboratively by small groups of students	Idem	Idem
Self-assessment tools, with criteria at both the individual and the group levels, will be finalised outside the seminar by the teachers	One framework for the self-assessment tools in all classes – criteria to be defined interactively with the students, (ensure that criteria make sense to students) – focus on assessment at the group level (not items at an individual level) – an open-ended rubric for “comments” is added	The assessment tool should include academic objectives (in addition to group functioning), while remaining focused at the group level
	Principles – carry out whole-class discussions with the students – work with them on handling possible disagreements within the group	Principles Continue to carry out whole-class discussions linked to assessment – to construct the criteria – to ensure reflection following assessment experiences
<b>Outside the seminar, on the teachers’ own initiative</b> The teachers working at the same grade level developed a single tool with the same assessment criteria; the tools differed between the grades (same criteria for individual and group levels in grades 3–4, different criteria in grades 5–6)	<b>Outside the seminar, on the teachers’ own initiative</b> For each of the criteria in the assessment tool negotiated with the class, the appraisal is communicated by open-ended comments written by the group	<b>After the three seminars</b> Additional meetings were held between the teachers and the researcher to define the classroom observations and the interviews to be conducted for longer-term research purposes
Use of the assessment tool immediately after the academic activity carried out by small groups of students		
Classroom observation: some teachers initiated a whole-class discussion about the assessments the students had carried out	Classroom observation: the teachers adopted the role of moderator when defining the criteria with the students, sometimes reformulating proposals and clarifying/regrouping certain proposals	

*Note* Both researchers (authors of this chapter) participated in the three seminars, but the classroom observations and post-seminar meetings were conducted by the second author in the context of his doctoral research

## References

- Allal, L., & Mottier Lopez, L. (2005). Formative assessment of learning: A review of publications in French. In OCDE (Ed.), *Formative assessment: Improving learning in secondary classrooms* (pp. 241–264). Paris: OECD.
- Allal, L., & Mottier Lopez, L. (2014). Teachers' professional judgment in the context of collaborative assessment practice. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 151–165). Dordrecht: Springer.
- Bergold, J., & Thomas, S. (2012). Participatory research methods: A Methodological approach in motion [110 paragraphs]. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 13(1). <http://www.qualitative-research.net/index.php/fqs/article/view/1801/3334>. Accessed January 22, 2015.
- Bourassa, M., Bélair, L. M., & Chevalier, J. (2007). Les outils de la recherche participative. *Éducation et Francophonie*, 23(2), 1–11.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32–42.
- Desgagné, S. (1997). Le concept de recherche collaborative: L'idée d'un rapprochement entre chercheurs universitaires et praticiens enseignants. *Revue des Sciences de l'Éducation*, 23(2), 371–393.
- Engeström, Y., Engeström, R., & Kärkkäinen, M. (1995). Polycontextuality and boundary crossing in expert cognition: Learning and problem solving in complex work activities. *Learning and Instruction*, 5(4), 319–336.
- Gilliéron Giroud, P., & Ntamakiliro, L. (Eds.). (2010). *Réformer l'évaluation scolaire: Mission impossible?*. Berne: Peter Lang.
- Gosselin, M., Viau-Guay, A., & Bourassa, B. (2014). Le développement professionnel dans une perspective constructiviste ou socioconstructiviste: une compréhension conceptuelle pour des implications pratiques. *Perspectives interdisciplinaires sur le Travail et la Santé*, 16(3). <http://pistes.revues.org/4009>. Accessed July 19, 2015.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Lussi Borer, V., & Muller, A. (2014). Quel apport/usage du “voir” pour le “faire” en formation des enseignants du secondaire. In L. Paquay, P. Perrenoud, M. Altet, J. Desjardins, & R. Etienne (Eds.), *Travail réel des enseignants et formation: Quelle référence au travail des enseignants dans les objectifs, les dispositifs et les pratiques?* (pp. 65–78). Bruxelles: De Boeck.
- Morales Villabona, F. (2013). *Modélisation d'une évaluation collaborative des apprentissages des élèves: Alternance entre un dispositif de formation et des pratiques en classe à l'école primaire*. Canevas de thèse en Sciences de l'éducation, Université de Genève, Suisse.
- Mottier Lopez, L. (2015). L'évaluation formative des apprentissages des élèves: Entre innovations, échecs et possibles renouvelés par des recherches participatives. *Questions vives*, 23. <https://questionsvives.revues.org/1692>. Accessed April 7, 2016.
- Mottier Lopez, L., Borloz, S., Grimm, K., Gros, B., Herbert, C., Methenitis, J., et al. (2010). Les interactions de la régulation entre l'enseignant et ses élèves: Expérience d'une recherche collaborative. In L. Lafortune, S. Fréchette, N. Sorin, P. A. Doudin, & O. Albanese (Eds.), *Approches affectives, métacognitives et cognitives de la compréhension* (pp. 33–50). Québec: Presses de l'Université du Québec.
- Mottier Lopez, L., Tessaro, W., Dechamboux, L., & Morales Villabona, F. (2012). La modération sociale: Un dispositif soutenant l'émergence de savoirs négociés sur l'évaluation certificative des apprentissages des élèves. *Questions vives*, 6, 159–175.
- Paillé, P., & Mucchielli, A. (2012). *L'analyse qualitative en sciences sociales*. Paris: Armand Colin.
- Race, P. (2001). *A briefing on self, peer and group assessment*. York: Learning and Teaching Support Network Generic Centre (Assessment series no. 9).

- Reed, E. S. (1996). *Encountering the world: Toward an ecological psychology*. New York: Oxford University Press.
- Salomon, G. (Ed.). (1993). *Distributed cognition: Psychological and educational considerations*. Cambridge, UK: Cambridge University Press.
- Schneuwly, B., Dolz-Mestre, J., & Ronveaux, C. (2006). Le synopsis: Un outil pour analyser les objets enseignés. In M.-J. Perrin-Glorian & Y. Reuter (Eds.), *Les méthodes de recherche en didactiques: Actes du premier séminaire international sur les méthodes de recherches en didactiques de juin 2005* (pp. 175–189). Villeneuve d'Ascq: Presses universitaires du Septentrion.
- Vinatier, I., Filliettaz, L., & Kahn, S (Eds.). (2012). Enjeux, formes et rôles des processus collaboratifs entre chercheurs et professionnels de la formation: Pour quelle efficacité. *Travail et Apprentissage*, 9.
- Voigt, J. (1985). Pattern and routines in classroom interaction. *Recherches en Didactique des Mathématiques*, 6(1), 69–118.

# Chapter 11

## Cooperative Learning About Assessment for Learning

Kari Smith

**Abstract** Assessment for learning (AfL) is salient in the rhetoric of policymakers and national steering documents in many countries. It has also been embraced by educators internationally, including those in Norway. However, despite the explicit positive intentions of all parties, there are many challenges in the process of AfL implementation in schools. One major challenge is the increased testing regime practised at a national level which presents teachers with the dilemma of whether to teach to the test to ensure high test scores or support learners in developing sustainable learning strategies. In other words, teachers often must choose between short-term and long-term learning effects. A second challenge is the isolation some teachers feel when practising AfL because it has not been integrated into the wider culture of their workplace; consequently, AfL is not valued equally by colleagues and school leadership. There is insufficient cooperation regarding AfL, and the stakeholders involved do not share a common assessment language. The current chapter presents a Norwegian project in which various stakeholders engaged in cooperative learning about assessment, focusing on the professional development in AfL of head teachers who lead subject matter teams. The project involved cooperation between a regional educational authority, an expert in assessment from the university (the researcher), the school principals, and four head teachers from each of five schools. The project's aim was to support head teachers' cooperative learning about how to improve assessment practice and to allow them to serve as agents of AfL in their own subject matter teams as well as in other teams in their schools.

### 11.1 Introduction

I was told by the principal to attend a half-day seminar on assessment for learning (AfL) given by a “big name” from the university. The seminar was good, and I became

---

K. Smith (✉)

Norwegian University of Science and Technology, Trondheim, Norway  
e-mail: kari.smith@plu.ntnu.no

convinced that assessment is important for students' learning in terms of motivation and helping them believe in themselves as learners. But I was the only teacher from my school attending the seminar, and there was no follow up later on. How can it be expected that I shall make changes by myself when the assessment system in school is mainly based on testing to prepare the students for the final exam? Besides, I do not know enough about how to practise AfL even though there is a leaflet with a list of techniques published by the Ministry, and the seminar leader gave some really good examples. But I do not understand why the different techniques are helpful; they are just items on a list.

I need time to really learn about AfL, the theory behind it, as well as the possibility to develop my own ways of doing it and trying it out in my own classes. A four-hour seminar is not enough for teachers to make changes.

(Norwegian secondary school teacher)

This comment and others similar to it were expressed by multiple teachers encountered by the representative of a regional educational authority and by myself as a university researcher in the area of assessment. The teachers' comments guided us in the planning and administration of a yearlong professional learning project for secondary school teachers focusing on assessment for learning (AfL). AfL has been on the political education agenda in Norway for several years; however, despite multiple initiatives, problems with implementation have remained, and the changes in classroom practice have not gone as expected.

This chapter first discusses the challenges of developing awareness of the pedagogical aspects of AfL and then elaborates on national initiatives to introduce AfL in Norway. Following this, the chapter outlines a joint initiative by a regional educational authority, a university representative, and five schools aimed at changing assessment in the respective schools. The initiative had multiple aims. First, we sought to inform school leaders that AfL practice would not occur unless they promoted its pedagogical value and supported staff development in AfL. Second, we proposed that subject head teachers be supported in changing assessment practices through action research, which would hopefully, with the support of the school principals, initiate a process towards a broader aim of developing an AfL culture in their schools. The length of the project did not, however, enable follow-up studies examining the extent to which the broader aim was achieved. The chapter ends with recommendations on how to avoid top-down models of change by investing in multiple small-scale, long-term development projects.

## 11.2 Pedagogical Aspects of AfL

In Norway, as in many other countries, there is a strong political focus on AfL, evidenced by steering documents and national funding for teachers' learning of AfL. However, the national efforts do not appear to have achieved the intended effects (Hopfenbeck et al. 2013). That said, efforts to introduce AfL in schools



should not be attributed to political decisions only. This perception may lead to AfL being seen as superficial and as only a set of techniques that teachers are instructed to introduce into their teaching. The pedagogical aspects of AfL are more likely to explain why AfL has been embraced by educators internationally, as well as in Norway. For deeper and sustainable changes to occur, teachers must be convinced of the pedagogical benefits of the changes and that the changes will be best for their students (Timperley et al. 2007). Hattie and Timperley (2007) relate AfL to feedback, which provides answers to students about the following questions:

Where am I going? (What are the goals?), How am I doing? (What progress is being made toward the goal?), and Where to next? (What activities need to be undertaken to make better progress?). These questions correspond to notions of feed up, feed back, and feed forward. How effectively answers to these questions serve to reduce the gap is partly dependent on the level at which the feedback operates. These include the level of task performance, the level of process of understanding how to do a task, the regulatory or metacognitive process level, and/or the self or personal level (unrelated to the specifics of the task). (Hattie and Timperley 2007, p. 86)

AfL processes are primarily evident in the manner in which teachers provide feedback to the students and how students make use of the feedback given. Hattie and Timperley (2007) suggest that feedback and AfL should not examine learning outcomes solely as achievements measured by grades. Rather, to improve achievements, other aspects of learning must be strengthened, such as self-efficacy and self-regulated learning processes. Learning will not occur if students do not believe in their own competence to learn and handle the learning tasks with which they are challenged. Bandura (1977, 1986) has indicated that the motivational aspects of learning can be understood by referring to a person's individual system of beliefs about her or his capacity to learn. The learner's internal feedback and individual capacity beliefs engage in a dialogue with external feedback from significant others, most commonly teachers and peers (Nicol and Macfarlane-Dick 2006). Useful feedback cannot be a monologue transmitted from the teacher to the student; the student must be an active partner in the dialogue, which can also engage peers. The quality and type of feedback plays a central role in developing learners' self-efficacy. Practising assessment for learning means that teachers are mindful and competent in their feedback on student learning.

Hattie and Timperley (2007) have shown that feedback is especially effective if it is given at a regulative and metacognitive level. AfL is about learning how to learn and developing self-regulated learning with both short-term and long-term perspectives. Zimmerman (1990) claims that self-regulated learners have better academic achievement and that 'self-regulated learners select and use self-regulated learning processes to achieve desired academic outcomes on the basis of feedback about learning effectiveness and skill' (pp. 6–7). AfL includes the understanding that assessment activities become genuine learning activities (Smith, in press). Hayward (2015) favours dropping the preposition in 'assessment *for* learning' in order that assessment and learning be so tightly linked in the curriculum that assessment becomes learning. However, she draws attention to challenges related to

how assessment is perceived by all stakeholders in education, perceptions also observed in Norway.

### 11.3 Challenges in Implementing AfL

In the literature, there is sufficient documentation about the pedagogical value of AfL. I would argue that for successful and sustained implementation of AfL, teachers must be shown not only the techniques for implementation, but must also develop comprehensive understanding of how and why AfL can promote learning. Techniques for implementation have become a focal point in the rhetoric of policymakers and steering documents in Norway. Despite the explicit positive intentions of these declarations, there are many challenges to the process of AfL implementation in schools (Hopfenbeck et al. 2013, 2015). A major challenge is the increased testing regime practised at a national level which presents teachers with the dilemma of whether to teach to the test to ensure high test scores (Popham 2001) or support learners in developing sustainable learning strategies. In other words, teachers commonly must choose between short-term and long-term learning effects (Hayward 2015; Smith 2011).

A second challenge is the isolation some teachers feel when practicing AfL because it has not been integrated into their school's culture, and consequently, it is not equally valued amongst their colleagues and the school leadership. Commonly, there is little cooperation regarding AfL. Hopfenbeck, together with colleagues from Oxford University and the University of Bergen, carried out a study of the implementation of AfL in municipalities and schools that participated in an extensive AfL project initiated by the Norwegian government (Hopfenbeck et al. 2013). The Assessment for Learning programme was launched in 2010, and the involved schools worked closely with the Norwegian Directory of Education and Training (DET) over a period of 16 months. This AfL initiative was a continuation of a previous programme (Improved Assessment Practice). 'The overall goal was to improve formative assessment practices in the classroom by developing distinct criteria to clarify how to reach curriculum goals' (Hopfenbeck et al. 2013, p. 28). This programme resulted in teachers devoting time to develop goals and corresponding criteria for every subject at every age level, without considering how to strengthen student learning to achieve these goals. Due to the evaluation of the first programme, policymakers decided to systematically implement AfL at a national level in Norway, and four basic principles for assessment became the core pillars of the project.

Students learn better when they:

1. Understand what to learn and what is expected of them.
2. Obtain feedback that provides information on the quality of their work or performance.
3. Are provided advice on how to improve.

4. Are involved in their own learning process and in self-assessment. (Hopfenbeck et al. 2013, p. 28)

Hopfenbeck et al. (2013) examined the implementation of the Norwegian AfL programme and conducted individual and group interviews with school leaders, teachers, and researchers. A main finding was that trust among the stakeholders, school leaders, and teachers was the key to success. Success stories were based on self-reports of the experiences of the interviewed stakeholders. However, no significant differences in learner achievements were observed between participating and nonparticipating schools. The research team concluded that not all schools achieved the expected goals, which were to change teaching practice and student involvement in assessment. In interviews with Norwegian professors who have expertise in AfL, Hopfenbeck et al. (2013) found that many of these professors criticized AfL implementation in Norway for the following reasons:

1. The variation among schools had not been sufficiently considered. A one-size-fits-all model does not work.
2. The national initiatives were not built on what teachers and school leaders find most challenging but on what the government finds challenging.
3. Researchers appear to emphasise the complexity of AfL, whereas the national programme indicated that there is a 'right way' of practising AfL that can be prescribed to teachers.

In Norway, many schools collaborate with researchers from the university, and teachers and principals likely feel tension between the researchers' message and the simplified version of AfL presented by the DET. In the transition between rationale and implementation, AfL policy has been reduced to techniques and ideas presented on a national website and in handbooks for teachers.

From the researchers' perspective, the concern was that the implementation of AfL has been superficial with no real understanding. Instead of acknowledging the complexity of the field, the DET offered a teacher friendly programme which could be implemented step-by-step, like a recipe. Even if the members of the DET strongly emphasised that these practices could be developed in several ways, some school leaders and teachers have interpreted the website version of AfL as the "truth" about AfL. (Hopfenbeck et al. 2013, p. 61)

One conclusion of the research team was that 'it is evident that there are challenges in how to transform the complex knowledge researchers possess into knowledge that teachers can use in developing their assessment literacy' (Hopfenbeck et al. 2013, p. 62).

## 11.4 Cooperative Learning About AfL

The work of Hopfenbeck et al. (2013) indicates the importance of developing a culture of AfL which, within a given framework, is unique to each school. The principal, teachers, and learners must develop a shared language of AfL that enables

individual practices (Smith and Engelsen 2012). Developing a culture and shared language relates more to work-based learning and less to off-job learning. Creating communities of learning characterized by trust, openness, and critical reflection on one's own practice has been found to be supportive for introducing changes to the school's and individual teacher's assessment practices (Smith and Engelsen 2012). Work-based learning within a community of practice has been advocated by many researchers claiming that professional learning occurs best within a supportive sociocultural learning environment (Eraut 2004, 2014; McNamara et al. 2014; Wenger 1998).

'A culture is a powerful, latent, and often unconscious set of forces that determine both individual and collective behaviour, ways of perceiving, thought patterns, and values .... Cultural elements determine strategy, goals and modes of operating' (Schein 1999, p. 14). A culture of assessment in school means that school leadership, teachers, and students have a shared understanding of assessment, assessment purposes, and how to practise assessment. A culture of assessment does not suggest that all practices are the same but that all practices are supported by a shared set of attitudes and values, with the explicit goal to promote learning. It means that students, teachers, leadership, and school authorities have developed a shared language when discussing assessment. In other words, the actors have become assessment literate, which according to Stiggins (1995), can be defined as follows:

Assessment literates know the difference between sound and unsound assessment. They are not intimidated by the sometimes mysterious and always daunting technical world of assessment (p. 240).

The aim of professional learning about AfL in the project presented here was to provide teachers with an understanding of the pedagogy inherent to AfL and with skills in translating this understanding into practice. The professional learning of teachers in AfL has recently been the focus of several Norwegian researchers in response to messages from central authorities. Most studies have examined how teachers develop individual AfL literacy (Engelsen and Smith 2014; Rønsen and Smith 2013), with little focus on staff development. In contrast, Engelsen and Smith (2014) describe how a three-year project successfully supported a Norwegian elementary school in developing a sustained AfL culture.

## 11.5 An Intervention in Support of AfL

The remaining sections of this chapter present a Norwegian project that involved a representative from the regional educational authority, a representative from the university, and school principals from five schools. They planned an intervention project with four head teachers from each school aiming to develop a shared understanding of and competence in AfL. A further aim was to enable the head teachers to empower their teachers in AfL in their respective schools. A more

implicit but important aim was that all stakeholders should become assessment literate so that AfL would be supported at a system level. The key elements of the project were introducing participants to the research literature on AfL and involving them in school-based action research so respective school teams could contextualise professional learning in their own schools and teaching practices. The specific question that the empirical section of this chapter seeks to answer is how the various actors in the project perceived the intervention as a process of individual and collective empowerment.

### ***11.5.1 Context of the Study***

The project began with cooperation between the local authorities and the university in 2009. The first aim was to examine teachers' and students' perceptions of assessment and the extent to which these two central assessment stakeholders used the same assessment language. The findings of this study suggested a wide gap between teachers' and students' understanding of assessment and the manner in which they viewed assessment practices. Specifically, this study revealed a lack of AfL competence among teachers (Havnes et al. 2012). Acting upon these findings, a professional research and development project was started under the responsibility of the local authorities and the university. Our understanding was that to implement AfL in schools, teachers and all stakeholders must be empowered in AfL practice. The current chapter focuses on data collected from the second cohort; 20 participating head teachers from five different schools worked on AfL with their school teams between the monthly course meetings. Moreover, three of the principals from the five participating schools participated occasionally. The group met with an external expert (the researcher) and the head of education at the regional authority on 7 days (8 h per day) over the school year.

### ***11.5.2 Intervention***

The intervention course had three main components: (1) sharing of current assessment practices, questions, and challenges faced; (2) interactive lectures on the pedagogical rationale underpinning AfL; and (3) team reading and presentation of the research literature on AfL and its uses in their own practice. The practical professional development tool in the intervention was action research (McNiff 2013), and the 'reflective circle' (Schön 1983) framed the three components. The participants were required to clarify their own point of departure at the beginning of the course and then formulated questions and areas for further education development. They were also required to learn about AfL through lectures and reading of relevant literature. Each team developed an action research (AR) plan that carefully documented AfL implementation. The concept of an action research project was

introduced to the participants at the very first meeting, and throughout the course, they could relate the content of the seminars to the various stages of the AR project. In the second meeting, they were asked to present the action research focus (question), and each team received suggestions for relevant research literature. In the third meeting, they presented the literature to the other participants and discussed how it related to their own project. Then they presented various stages of the project for feedback from the course leader (university professor) and their fellow participants. Upon completion of the course, all projects were presented at a two-day seminar for representatives from the regional school authorities, principals from all of the schools, and participants from the first cohort. The projects were also presented to all teachers in the respective schools.

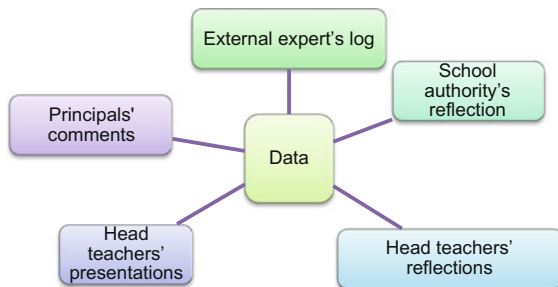
Previous experiences with similar professional learning models combining top-down aspects (requirement to engage in action research, reading the relevant literature suggested by the external expert, and lectures on AfL) and bottom-up approaches (participants selecting themes for their own focus of development, sharing experiences, and presenting relevant literature) have been found to have translational effects for assessment practice (Engelsen and Smith 2014; Smith 2011). Whereas previous projects concerned professional development activities with the whole staff in small schools, the current project strongly focused on action research as a professional development tool. A more formal presentation of the action research (AR) project was required because the head teachers would be mentoring other teachers in larger secondary schools. Documentation and articulation of their own professional development in relation to AfL was therefore viewed as important.

### ***11.5.3 Methodology***

To learn about the perspectives of the different actors, several qualitative data collection instruments were used, as presented in Fig. 11.1.

The expert wrote a reflective log after each session in which she recorded her impressions of the didactic aspects of the seminar, interactions with the group, her own feelings, questions, and doubts, and attempted to hypothesize how the participants felt. The representative from the regional authority who had been active in all seminars wrote a reflective note at the end of the project, focusing on her learning and to a larger extent how she felt the school teams were progressing throughout the project. The head teachers wrote continuous reflections, especially about their own learning processes, work by their team, and their worries in relation to AfL implementation. The challenges of conducting action research for the first time became a central theme. The final action research projects documented the outcome of the learning processes of the head teachers and how they planned to implement AfL in their own context. The participating principals wrote a few comments at the end of the course.

**Fig. 11.1** Data collection instruments



The analysis of the data as presented here is the subjective interpretation by the external expert; however, the content was presented to all participants as a PowerPoint presentation inviting comments. The discussion that followed the presentation suggested minor changes to the interpretations, one of which was related to the importance of reading and presenting the literature on AfL to each other. Moreover, the first full version of the chapter was approved by all participants.

### 11.5.4 Findings

The main findings are presented in relation to the various stakeholders and the processes they underwent.

The head teachers initially felt strongly the tension between directives, theoretical input, and external exams, which resulted in more general pressure for secondary school teachers.

It seems as if we teachers are pulled in three directions, and sometimes I feel as if I am the server of three masters who do not talk to each other, the final exams at the end of the year, all the documentation I have to do to avoid getting sued by students or parents, and all the pedagogy of how to support each and every student.

However, towards the end of the course, a more positive and future-oriented tone could be observed in reflective notes:

To understand my own practice and change it based on theoretical knowledge was new to me.

Previously, self-assessment was added at the end and after my assessment. Now, the students assess their own and their friends' processes and work in progress.

The head teachers moved from being critical of AfL and defensive of their own practice toward a more open-minded and positive attitude regarding the changes they had made.

In her log, the external expert mirrored the head teachers' changed attitudes. After the first session in September, she wrote: 'Same once more, a nice group, they know little about AfL and AR. They are very defensive of own practice.' In

November, however, the log had a more optimistic tone, as the school teams had worked with the literature and presented their understandings to the group. The external expert observed a change in the head teachers' attitude toward AfL and beginning understandings of the underlying pedagogical principles: 'Good discussions of articles, beginning of change of attitude. I wonder what their AR projects will be about....'

After the Christmas break in February, the teams had started to work on their AR projects, and the external expert became aware that the scope of the projects was too wide and that the teams would be unable to make all of the desired changes, especially when goals involved creating deep and sustainable changes: 'They want to change everything. I need to get them to focus their projects. They need to learn how to work with AR as a tool for development and change, which they can continue to use at the end of the course.'

Finally, when the presentation seminar occurred in June, the outcome of the participants' learning was presented to a wide and important audience. The external expert wrote: 'I feel like a proud teacher at graduation. The projects are good; they have really found their own interpretations of AfL in their own school context. I am truly pleased.'

The representative from the school authority, an experienced teacher, had been following the first cohort and was well acquainted with AfL from steering documents and her own prior experience as a teacher. She wrote her reflections at the end of the course, acknowledging the process the head teachers had undergone. She felt that the head teachers first focused on student learning separate from assessment before developing an awareness of the teacher's role. Teachers realised that the aim of AfL was not to document all assessment practices to guard against being sued by parents and students. Instead, AfL was about adapting assessment to the class context and to discussions in meetings with individual students. At the end of the course, she observed that the head teachers made clear links between teaching and assessment and did not see these activities as separate and incompatible. Regarding her own learning in the project, she wrote:

It is important to mention the professional learning I have gained throughout the project period. I have learned about theoretical aspects of assessment and received a better understanding for the teachers' challenges in the classroom. The fact that some of the head teachers took the time to write a final paper provided a fascinating insight into the head teachers' learning process in the project.<sup>1</sup>

The school principals were asked to comment on the course at the end of the final seminar, and the following comment presents a shared opinion among the school leaders: 'As a principal and teacher, I have a new understanding of assessment. The school will continue to work on AfL and dedicate discussion and development time for all teachers to be involved.'

---

<sup>1</sup>The presentation of the project was compulsory, and writing a final paper was voluntary.



A declared intention to develop an AfL culture in school appears to be in place but does not necessarily guarantee that there will be follow up. The current study does not inform about the project's sustainability.

The most promising findings can be found in the action research presentations of the head teachers. The topics that they decided to work on with their teams varied, as the following examples indicate:

1. Feedback in physical education (*school specialising in sport*)
2. Implementation of Black and Williams's (2009) five principles for AfL:
  - (a) Develop goals and criteria for success with students
  - (b) Create activities that represent students' understanding
  - (c) Provide clear and useful feedback/feed forward
  - (d) Create situations in which students support each other
  - (e) Enable students to create ownership of their own learning  
(*civic studies teacher team*)
3. Peer and self-assessment in language teaching to promote students' self-confidence, motivation, and reflection (*team of language teachers*)
4. How to develop a shared understanding of grades concerning order and conduct?<sup>2</sup> (*interdisciplinary team*)
5. Assessing the learning of mathematics using less tests (*math teachers*).

The projects reflected the concern of the participants within their own teaching context and presented documentation of changed assessment practices supported by theory, as well as ideas of how to make future changes. The great variety in the topics created mutual learning at the school level, as head teachers developed practical ideas for change processes. Further, AfL practices in multiple school subjects were presented and made available to all schools in the region through postings on the internal website for all regional secondary schools.

## 11.6 Discussion

In this discussion, some key issues from the findings will be addressed to serve as a framework for developing an AfL culture in schools and provide teachers with ownership of changes. In other words, this discussion will be directed towards suggestions for empowering teachers to becoming independent and confident practitioners of AfL.

Many educators, such as the secondary school head teachers and principals in this study, work within an educational system under accountability pressure and a widely developed testing regime (Darling-Hammond and Snyder 2015; Hayward

---

<sup>2</sup>In Norway, students are given a grade in 'order and conduct,' reflecting the students' punctuality, behaviour, bringing necessary equipment to lessons, doing homework, etc.

2015). In some contexts, such as Norway, AfL is high on the educational political agenda, and extensive rhetoric addresses the importance of AfL as a method to improve students' test scores (Engelsen and Smith 2014; Hopfenbeck et al. 2015). Less attention is given to the pedagogies that underpin AfL because AfL is considered as a golden key to achieve learning outcomes that are compatible with expectations at the national and international levels (Smith, in press). However, the core of AfL is the communication between the teacher and learner about the progress of learning related to a specific task, specific subject, or school. This pedagogical focus of AfL has not been given sufficient attention in teachers' professional development activities. Teachers are the ones who care for the individual learner and how he or she progresses. The secondary school head teachers in this study felt caught between concern for the individual student and the need to ensure that external demands, such as high achievement scores, are met. The decisions that teachers make are formed by the context, situation, class, and individual student. They must find a balance between external demands and pedagogical considerations. When the external pressure is too high, teachers are more likely to 'teach to the test,' which might lead to better exam scores in the short term but not necessarily to more meaningful learning. The ongoing struggle to balance meaningful learning and accountability pressures is not new and deserves more attention from the research community, policymakers, and practitioners. The findings in this study reveal the tension that secondary school head teachers especially must contend with. What supported the head teachers in developing an individual understanding of how to practise AfL and promote it to their own teacher teams was the emerging theoretical understanding of how AfL impacts student motivation and self-regulation, and their trust in their own competence to learn. The head teachers were familiar with techniques and regulations issued by the Ministry regarding how to practise AfL, but they had little or no understanding of the pedagogical basis for the imposed changes. By reading the suggested literature, head teachers became able to link AfL to personal pedagogical values.

Another central issue regarding our findings is that teachers are used to being told what to do, such as how to implement AfL. We know less about how they learn about AfL and how they develop assessment literacy (Engelsen and Smith 2010; Hayward 2015; Rønsen and Smith 2013; Smith 2011). Ample research has indicated that top-down approaches do not work well; a good balance between bottom-up processes (empowerment) and the parameters of a given framework (top-down guidance) appears to be a more effective approach (Engelsen and Smith 2010, 2014).

This one-year project was too short to develop sustainable changes; longer projects that implement continuous support during change processes are more likely to ensure sustainability (Engelsen and Smith 2014; Timperley et al. 2007). The development stages that head teachers in the current study underwent during the course resonate with the phases Rønsen (2015) detected in her doctoral work, which she called 'the preparatory phase,' 'the theorizing phase,' and 'the explorative phase.' At the end of the preparatory phase, teachers in Rønsen's study were able to see their own practice through reflective discussions within a community of

learning, which included other teachers and external experts. The teachers went from being defensive of their own practice to being able to reflect on their own practice with a critical perspective. In the second phase, the theorizing phase, teachers developed a shared language of assessment by reading the relevant research literature and thus became able to articulate and discuss their practices using the assessment language. This phase helped teachers develop a meta-perspective on their own teaching. In the final ‘explorative’ phase (after nearly three years), teachers developed an individual practice theory which enabled independent AfL practice, and they were able to support their actions with theory (Rønsen 2015).

Two primary conclusions can be drawn from the current small study and the more in-depth study of Rønsen. First, the starting point for change should be the teachers’ current assessment activities. They must critically reflect on their own practice to develop an understanding of how assessment impacts student learning. The process of self-examination is facilitated by collegial discussions within safe communities of learning. Second, the support of external expertise in the form of discussions based on the relevant research literature strengthens the teachers’ learning processes in developing a personal practice theory of assessment, which allows them to become assessment literate (Engelsen and Smith 2014).

However, changes throughout an entire school will not develop unless the assessment culture in the school is changed. To do so, not only teachers but also the leadership of the school should be involved (Hill 2011; Leithwood et al. 2004; Printy 2008; Smith and Engelsen 2012). In the study reported in this chapter, the school leadership was invited to join the project from the beginning. The five principals were involved in planning the intervention. Some principals participated in the intervention, whereas others only attended the final seminar for project presentations. Those who participated in most meetings developed a shared language with their head teachers when discussing assessment, and more collective learning about AfL occurred in these schools, which will likely impact future AfL developments in the schools. The involved head teachers were part of the pedagogical and didactical school leadership and in the position to initiate change processes in their respective schools. The data were collected during and at the end of the project period (intervention), and no follow-up data have been collected. Thus, we do not know if the intentions expressed by the leadership were put into practice, a process which will require time, resources, and a strong environment of trust in order to enable a variety of practices and tests of new ideas. As Engelsen and Smith’s work in another project has shown (Engelsen and Smith 2010, 2014; Smith and Engelsen 2012), the principal was clearly the driving force for creating a sustainable AfL culture in the school.

The final point to be highlighted in this discussion is the use of action research as a professional development tool. Previous research has documented successful use of action research as a tool to strengthen teachers’ professional learning, the main advantage being the ownership that practitioners develop in association with their own learning (Kane and Chimwayange 2014; McNiff 2013; Smith and Sela 2005). However, it cannot be expected that teachers will engage in action research projects

unless they are given time and resources, as well as careful and patient guidance from experienced researchers. Ponte et al. (2004) argue that talking to practitioners about action research and how it is conducted does not help and confuses individuals unfamiliar with research, particularly practitioner research. Practitioners obtain the sense of action research and become aware of its meaningfulness only when they start working on their own projects.

Action research involves understanding one's own practice, learning about the specific issues, testing alternatives through systematic documentation, which is analysed, and interpreting results and conclusions. The manner in which action research was used in the current study reflected the well-known cycle or spiral of experiential learning (Kolb 1984), as well as Korthagen's (1985) ALACT model. The basic idea is that through systematic critical analysis of current assessment practice, head teachers seek new alternatives, which are based on new knowledge acquired through collective learning about AfL (see the description of the intervention above). In the current project, action research was conducted in teams, not by individual teachers. Team projects, which enable teachers to share the workload and provide time and space for discussions, exchanges of experiences, and sharing of responsibilities, are less frightening to teachers than individual projects. Change processes become a joint venture and not an individual process without opportunities for peer dialogue, and there is a constant flow of ideas within the research groups. The voice of the teacher presented at the very beginning of the chapter supports this argument.

## 11.7 Recommendations

In this project, no golden key was discovered for developing teachers' AfL competence. However, this project provided evidence of head teachers developing their own personal, team, and school practice theories of AfL. For example, in one school, teachers were unhappy with the formal grades given to students for order and conduct; the grades were primarily based on the number of reprimands documented, and each individual teacher generally decided what should be documented. The four head teachers from this school decided to work together and conduct action research on this important cross-disciplinary issue. These head teachers involved all teachers in the school and the students in developing a school code of conduct. The conduct and order grades thus became transparent and closely linked to the school's code of conduct. Another example in which the intervention and action research changed assessment practice could be seen in how the head teacher of mathematics in one school, together with her team of teachers, developed and tested supplementary assessment tools. They introduced math portfolios with student-selected entries, and they invited students to write test items and be involved in correcting their peers' and their own tests. They introduced group testing in mathematics. These were huge changes for a generally traditional group of math teachers who were empowered by the goal of developing AfL.

The intervention project, in all successive cohorts, found that detailed top-down directives indicating how to practise AfL do not help teachers. All actors involved with school changes for improvement need to become assessment literate; thus school leaders, teachers, and school authorities must be involved in the same learning processes to develop a shared understanding of what creating an AfL culture means. The changes presented in the examples above could not have occurred if the regional authority had not invested money in the project and followed it closely in order to accept and understand the changes that occurred in the schools they controlled. Moreover, the school leadership, represented by the principal and head teachers, invested in their own learning of AfL and therefore could contextualise the changes in their own school. They were empowered to act as brokers of AfL within their respective schools. The responsibility for change was not left to individual teachers. The various actors were empowered in practicing AfL, which requires a shared language, individual and collective competence, autonomy, and responsibility. Control and uniform detailed directions from policymakers have not proven to be successful in AfL implementation in Norwegian classrooms. The general principles of the intervention presented here could serve as an example of how cooperation between various stakeholders in education can lead to changes within a given framework. However, AfL implementation requires openness to the specificity of schools, subjects, and teachers; in other words, one size does not fit all.

These results must be interpreted cautiously; no generalization of this small and limited study can be made, and a direct transfer of the model to other contexts should be avoided. Each context is unique. However, when changes are imposed on teachers in a top-down manner, they likely produce only cosmetic results. Teachers will accept deeper changes only if they address their main concern: namely, to support student learning (Day et al. 2005). Thus, models of change that seek to develop not only a shared understanding of the change but also an ownership of the changes (bottom up) and autonomy to adapt changes to suit personal practices (Timperley et al. 2007) are more likely to be effective. Our findings indicate the importance of developing a theoretical and pedagogical understanding of AfL rather than focusing only on its practical and technical aspects. The theoretical and research literature, which underpins the more practical aspects of AfL, needs to be shared with practitioners by engaging them in cooperative learning; this means involving them in reading, presenting, and discussing the literature in practice-based communities.

Finally, a main recommendation from this chapter is for policymakers and leaders of education to invest in multiple small-scale, long-term projects instead of multiple large-scale, short-term activities. The latter remains unfortunately the most common form of implementation of educational change.

## References

- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191–215.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Education Assessment, Evaluation and Accountability*, 21, 5–31.
- Darling-Hammond, L., & Snyder, J. (2015). Meaningful learning in a new paradigm for educational accountability: An introduction. *Education Policy Analysis Archives*, 23(7), 1–4.
- Day, C., Elliott, B., & Kingston, A. (2005). Reform, standards and teacher identity: Challenges of sustaining commitment. *Teaching and Teacher Education*, 21(5), 563–577.
- Engelsen, K. S., & Smith, K. (2010). Is excellent good enough? *Education Inquiry*, 4. <http://www.education-inquiry.net/index.php/edui/article/view/21954>. Accessed September 10, 2011.
- Engelsen, K. S., & Smith, K. (2014). Assessment literacy. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 91–107). Dordrecht: Springer.
- Eraut, M. (2004). Informal learning in the workplace. *Studies in Continuing Education*, 26(2), 247–273.
- Eraut, M. (2014). Developing knowledge of qualified professionals. In O. McNamara, J. Murray, & M. Jones (Eds.), *Workplace learning in teacher education* (pp. 47–72). Dordrecht: Springer.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Havnes, A., Smith, K., Dysthe, O., & Ludvigsen, K. (2012). Formative assessment and feedback: Making learning visible. *Studies in Educational Evaluation*, 38, 21–27.
- Hayward, L. (2015). Assessment is learning: The preposition vanishes. *Assessment in Education: Principles, Policy & Practice*, 22(1), 27–43.
- Hill, M. F. (2011). ‘Getting traction’: Enablers and barriers to implementing assessment for learning in secondary schools. *Assessment in Education: Principles, Policy and Practice*, 18(4), 347–364.
- Hopfenbeck, T., Tolo, A., Florez, T., & El Masri, Y. (2013). *Balancing trust and accountability? The Assessment for learning programme in Norway*. Paris: OECD.
- Hopfenbeck, T. N., Florez Petour, M. T., & Tolo, A. (2015). Balancing tensions in educational policy reforms: Large-scale implementation of Assessment for Learning in Norway. *Assessment in Education: Principles, Policy & Practice*, 22(1), 44–60.
- Kane, R., & Chimwayange, C. (2014). Teacher action research and student voice: Making sense of learning in secondary school. *Action Research*, 12(1), 52–77.
- Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice Hall.
- Korthagen, F. A. J. (1985). Reflective teaching and preservice teacher education in the Netherlands. *Journal of Teacher Education*, 9(3), 317–326.
- Leithwood, K., Seashore Louis, K., Anderson, S., & Wahlstrom, K. (2004). *Review of research: How leadership influences student learning*. New York: The Wallace Foundation.
- McNamara, O., Murray, J., & Jones, M. (Eds.). (2014). *Workplace learning in teacher education*. Dordrecht: Springer International.
- McNiff, J. (2013). *Action research: Principals and practice* (3rd ed.). Abingdon: Routledge.
- 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.
- Ponte, P., Beijard, D., & Ax, J. (2004). Don’t wait till the cows come home: Action research and initial teacher education in three different countries. *Teachers and Teaching: Theory and Practice*, 10(6), 591–621.
- Popham, W. J. (2001). Teaching to the test? *Educational Leadership*, 58(6), 16–20.

- Printy, S. M. (2008). Leadership for teacher learning: A community of practice perspective. *Educational Administration Quarterly*, 44, 187–226.
- Rønsen, A. K. (2015). *Vurdering som profesjonskompetanse (Assessment as professional competence)*. Unpublished doctoral dissertation, Faculty of Psychology, Bergen, University of Bergen, Norway.
- Rønsen, A. K. & Smith, K. (2013). Influencing and facilitating conditions for developing reflective assessment practice. *Professional Development in Education*, 1–17. doi:10.1080/19415257.2013.836126
- Schein, E. (1999). *The corporate culture survival guide: Sense and nonsense about culture change*. San Francisco, CA: Jossey-Bass.
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Smith, K. (2011). Professional development of teachers: A prerequisite for AfL to be successfully implemented in the classroom. *Studies in Educational Evaluation*, 37, 55–61.
- Smith, K. (in press). Assessment for Learning: A pedagogical tool. In D. Wise, L. Hayward & J. Pandya (Eds.), *The SAGE handbook of curriculum, pedagogy and assessment*. London: Sage.
- Smith, K., & Engelsen, K. S. (2012). Developing an assessment for learning culture in school: The voice of the principals. *International Journal of Leadership in Education*, 16(1), 106–125. doi:10.1080/13603124.2011.651161
- Smith, K., & Sela, O. (2005). Action Research as a bridge between pre-service teacher education and in-service professional development. *The European Journal of Teacher Education*, 28(3), 293–311.
- Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), 238–245.
- Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration*. Wellington, New Zealand: Ministry of Education. <http://educationcounts.edcentre.govt.nz/goto/BES>. Accessed May 1, 2015.
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge: Cambridge University Press.
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An Overview. *Educational Psychologist*, 25(1), 3–17.

# Chapter 12

## Developing Assessment for Learning Practice in a School Cluster: Primary and Secondary Teachers Learning Together

Sue Swaffield, Roszalina Rawi and Amanda O'Shea

**Abstract** The nature of professional development for the sustained implementation of assessment for learning (AfL) is a pressing and perennial challenge. So too is pupil transfer between schools. This chapter explores how cross-phase collaborative learning supports the development of AfL practice. Teachers from a secondary school and its seven feeder primary schools worked together using three principles of AfL derived from previous research to assist analysis of existing practices, to plan developments, and to scaffold discourse. The value of this approach was revealed by data gathered through questionnaires and interviews with teachers in the working group, along with observations of workshops at which participant teachers shared their work with other teachers, who were also invited to complete a questionnaire. It is suggested that AfL, underpinned by the principles of making learning explicit, promoting learning autonomy, and focusing on learning, can act as a pedagogical unifier across age ranges and subjects and thus aid pupil transfer. Recommendations are proposed for teachers, policymakers, and particularly school leaders.

### 12.1 Introduction

The challenge at the heart of the project discussed in this chapter is professional development for the sustained implementation of assessment for learning (AfL). Specifically, the research aimed at understanding how cross-phase collaborative learning between primary and secondary teachers supports the development of AfL practice. Pupils' learning, both in discrete subjects and in learning how to learn, occurs in many contexts with many different teachers over their school career.

---

S. Swaffield (✉) · R. Rawi  
University of Cambridge, Cambridge, UK  
e-mail: ses42@cam.ac.uk

A. O'Shea  
University of Northampton, Northampton, UK



Whilst AfL practices should be adapted to age ranges and subjects, they also need to be complementary and sustained so that learners experience coherence across settings and progression over time. Implementation is therefore not just about the practices of individual teachers but also about congruence throughout the primary and secondary schools children attend.

Coherence with appropriate adaptations to context can be achieved through a shared understanding of fundamental issues and adherence to common underlying principles. Swaffield (2011) argued that not all practices purported to be AfL support pupils in learning how to learn, because the practices are not in accordance with three key principles. These principles were formulated by a large multi-university team working with over 40 schools in England (James et al. 2007) and can be summarised as: making learning explicit, promoting learning autonomy, and focusing on learning rather than grades (James and Pedder 2006).

Application of AfL principles across a family of schools can be promoted by teachers from different schools and subject specialisms working together to improve their AfL understanding and practices. The work reported here involved collaboration among serving teachers, school leaders, and university researchers.

The next section draws on research literature to consider issues relating to teachers' collaborative professional learning and the development of classroom practice, the implementation of AfL, and pupils' transfer from primary to secondary school. This is followed by an overview of the particular project that provided the empirical data, and details of data gathering. Findings are then presented and discussed in advance of a conclusion that includes four themes arising from the cross-phase professional development work, an outline of some of the challenges for implementation of AfL, and recommendations for school leaders, teachers, and professional development policymakers.

## **12.2 Background Issues**

### ***12.2.1 Collaborative Professional Learning***

The importance of teachers' collaborative professional learning for improving the quality of teaching and learning in schools has been widely recognised, for example, by Borko (2004), Little (2005), Katz and Earl (2010), Pedder and Opfer (2011), and Willis et al. (2013). Teachers have long worked and learned together in groups variously termed 'professional learning communities,' 'collaborative learning communities' (Cooper and Boyd 1998), 'communities of practice' (Wenger 1998), and 'teacher learning communities' (William 2007a). In essence, while working to improve their practice, professionals learn together through interacting regularly, sharing resources, reflecting on practice, reviewing outcomes, and providing feedback and support for each other. Collaborative learning is grounded in practice, and those involved have been found to be more dedicated in

their efforts to improve teaching and learning (Rosenholtz 1989). Efforts to improve teaching practices should centre on peer collaboration and concerted discussions in teacher learning communities (Darling-Hammond 1996; Clausen et al. 2009; Wiliam 2007a). Clausen et al. (2009) pointed to the importance of a positive start and early success to motivate members and sustain their undertakings.

Putnam (2000) and Woolcock (2001) considered three forms of social capital—bonding, bridging, and linking. Strong support and empathy from colleagues who face similar challenges help teachers in learning communities build *bonding* social capital, while extended networking can develop *bridging* social capital (Putnam 2000), giving teachers access to valuable ideas and information. Wiliam (2007b) suggested that the most promising approach for implementing effective formative assessment practices is through teachers learning in small, building-based groups. These he contrasted with teachers meeting others from different schools, characterised as a good way to get new ideas about AfL, but described as ‘sources of information, not sources of change in teacher behaviour’ (Wiliam 2007b, p. 40). Smith (2011) suggested that a major disincentive to teachers’ professional development in AfL is the lack of competence of teacher educators in AfL. In order for AfL to be common assessment practice in schools, there needs to be improvement in processes at various levels of the education system. In a community of practice, novices can connect with experts, and academics with practitioners (Lave and Wenger 1991). Woolcock (2001) considers these vertical connections as *linking* social capital that allow teachers to leverage wider resources from beyond their usual peer groups.

### ***12.2.2 Teachers Innovating Practice in AfL***

Reviews and research, for example by Black and Wiliam (1998), Gardner (2006), and Earl (2012), indicated that AfL has the potential to make a substantial positive impact upon learning. However, classroom assessment practice can be extremely difficult to change. Partly, the difficulties lie in the translation of those theories into practice.

Wiliam (2009) contended that a useful model of development is to enable teachers to see what AfL means in practice and to understand what will make the greatest impact upon learning. Teachers need practical techniques for achieving the principles, modelled by practitioners in contexts that are relevant to them, alongside the flexibility to innovate with those techniques. Gardner (2010) argued that teacher self-agency is a key to change that is more powerful than notions of ‘theory-first or practice-first’ (p. 133). Teachers need to be supported in developing aspects of their practice as a personal agenda.

Sato et al. (2005) described two very different teachers developing AfL within their own contexts. This research and that of others (e.g. James et al. 2007) demonstrated that teachers’ practices are entwined with their beliefs and values. While teachers might try something new, it is likely to be discarded if it does not fit

within their personal modes and values of teaching. Sato et al. (2005) concluded that such changes do not come easily because they are essentially personal in nature. Implementing new ideas led nevertheless to changes in teachers' underlying beliefs and assumptions about teaching and learning. Long-term sustainable change arose from valuing continued reflection and experimentation with colleagues over time so that those innovations became internalised.

Suurtamm et al. (2010) similarly concluded that collaborative professional development (CPD) was instrumental in developing practice since it enabled teachers to network with others. Where such networks and communities have support, time, and space, they have proved successful in developing social capital with mutual trust and support, and intellectual capital with ideas and practice (James and McCormick 2009). Networks are an opportunity to develop bonding social capital between groups of teachers within similar contexts and bridging social capital between teachers of different key stages and subjects.

In terms of time, Wiliam (2009) argued that teacher learning communities should ideally run for at least two years, with groups of between eight and twelve teachers meeting approximately once a month being the most effective. Meetings should focus on allowing time for each teacher to feedback to the group on their classroom innovations but also include an element of new study to inform them about AfL and encourage future action planning.

### ***12.2.3 Issues with Implementing AfL***

One crucial element to implementing effective AfL lies in its conception and definition. However, while some teachers conceptualise AfL in terms of pupil autonomy, others focus on monitoring performance (Hargreaves 2005). James and Pedder (2006) argued that teachers do not necessarily practise what they most value in assessment, which sometimes leads to a focus on learner performance rather than autonomy. It was this research as part of the *Learning How to Learn* project that identified the three dimensions to AfL practices mentioned above: promoting pupil autonomy, making learning explicit, and focusing on learning. The difficulty for teachers was in balancing the need to monitor performance with other practices in assessment that might improve learning.

Tierney's (2006) review of influences on changing assessment practices highlighted time as a considerable factor. There needs to be the will and acceptance of the long time-frame associated with implementing reform when the impact on learning may not be immediately realised. Time is necessary for changes to become embedded within practice, yet there is often pressure to move on to new initiatives. Embedding change in assessment practices is not simple and requires support from school leadership (Swaffield and MacBeath 2006; Smith and Engelsen 2013). Kellard et al. (2008) cite the importance of support and enthusiasm from schools' senior leadership in terms of practical time and resources but also as enablers to a bottom-up approach towards change.

### **12.2.4 Transfer**

Galton et al. (2003) used the term *transfer* in the context of addressing research that involves pupil *transfer* between schools, rather than *transitions* between years within the same school. According to Mizelle (2005) an effective framework to facilitate such transfer should include a ‘vertical team’ (p. 59) that includes teachers and administrators across grade levels and schools discussing how to better align curriculum. Teacher intervention in the form of formative assessment activities that enhance self-esteem and motivation to learn can also ease the transfer issues faced by students (Craven et al. 1991; Ginsburg-Block et al. 2006; Miller and Lavin 2007). Beaumont et al. (2014) acknowledged the difficulty and challenges of transfer from secondary to higher education in respect of assessment practices and feedback.

Issues of pupil transfer between schools, innovating and implementing AfL, and teacher collaborative learning were all central to the project discussed in this chapter.

## **12.3 Project**

### **12.3.1 Background and Project Details**

The project involved a semi-rural secondary school (pupils aged between 11–16 years) and its seven feeder primary schools (pupils aged between 4–11 years), which together have approximately 130 teachers. The cluster has a long history of collaboration, and systemic developments in recent years (for example the creation of *Teaching Schools Alliances*) have seen additional opportunities and renewed commitment from senior leaders to enabling ongoing liaison among teachers. Teachers working together on projects with specific foci is one way in which the perennial aim of supporting pupils throughout their schooling is addressed.

The schools have close links with the local university, particularly through teacher education programmes (both initial and continuing), and research-focused partnerships. A senior leader of the secondary school with responsibility for cluster liaison secured a small amount of funding and approached the university to contribute to a cluster AfL project. In liaison with the school senior leader, the university colleague (first author of this chapter) facilitated professional development activities, fed in research, supported the teachers in elements of enquiry within the professional development activities, and researched the process.

The project, with its broad aim of raising attainment through AfL, was conceived at a cluster steering meeting as a way of building on the benefits of recent

partnership working among the schools. It lasted one academic year, starting and finishing with a joint staff meeting for all the teachers in the cluster. Senior school leaders (primary and secondary) addressed the joint staff meeting at the beginning of the year, explained the particular focus on feedback, and the plans for the whole year. These involved a working group of volunteer teachers drawn from all the schools meeting periodically, and a conference at the end of the year for all teachers to learn from the working group. The university colleague then invited the teachers to consider their AfL feedback practices in the light of pertinent research, in particular three principles of AfL practice (James and Pedder 2006), and outlined a framework for collaborative working. Subsequently, more teachers volunteered for the working group than could be accommodated, so the school leaders selected 23 (12 primary and 11 secondary) school teachers representing a variety of teaching experience, age ranges taught, and (in the secondary school) subject specialisms. The working group included the teaching head of one of the primary schools.

The working group, facilitated by the university colleague, met for two hours on each of four occasions, planning together, reporting progress, giving one other feedback, and considering published research (see Table 12.1).

The three AfL principles introduced at the joint staff meeting—making learning explicit, promoting learning autonomy, and focusing on learning rather than grades (James and Pedder 2006; James et al. 2007)—were referred to repeatedly and became part of the discourse providing a common language for discussing disparate practices. The five cross-phase subgroups formed around teachers' specific interests focusing on aspects of feedback in different contexts (as detailed in Table 12.2).

**Table 12.1** Full staff and working group meetings

Timeline	Type of meeting	Purpose
September 2013	Cluster full staff meeting	Introducing project and AfL principles
		Inviting volunteers for working group
October 2013	Working group meetings	Details of project purpose and process
		Identifying specific foci and forming subgroups
		Planning including initial data gathering from pupils/colleagues
November 2014		Findings and examples from international research
		Sharing practice, plans, and progress
February 2014		Subgroups reporting and feeding back to each other
April 2014		Sharing developments
		Planning for joint staff meeting
June 2014	Cluster full staff conference	Working group subgroups leading workshops, with explicit reference to three AfL principles
		Plenary discussion

**Table 12.2** Subgroups, practices/resources developed and principles

Subgroup: subject focus	Teachers/School phase	Name of resource/practice	Short description	AfL principle in focus
A: Mathematics	2 Primary	Topic targets	Continual self-assessment against targets with success criteria expressed simply and represented visually	Promoting learning autonomy Focus on learning
	2 Secondary	Focused feedback	Giving focused feedback to improve mathematics	Focus on learning
	3 Primary	Learning ladders	Progressive targets in pupil-friendly language, scaffolding pupils' discussion of their achievements and next steps	Making learning explicit Promoting learning autonomy
	2 Secondary			
C: Outdoor learning	2 Primary	Den building	Pupils work in teams to build a den, learning together in order to address success criteria	Making learning explicit Promoting learning autonomy
Physical education Music	1 Primary	Quality talk and oral feedback	During lessons pupils coach one another using shared success criteria	Promoting learning autonomy
	2 Secondary			
D: Science Geography	2 Primary	Grade boosters	Marked but ungraded work returned to pupils, with practice questions ('grade boosters') that target weaknesses	Focus on learning
	3 Secondary			
E: English (writing in history and modern foreign language)	2 Primary	Leave a line/TOWER model	Pupils leave a line between their writing to create space for peer feedback and improvements. TOWER model refers to: Talk, Organise, Write, Edit and Reflect	Promoting learning autonomy
	2 Secondary			

### ***12.3.2 Researching Cross-Phase Professional Development***

Data were generated towards the end of the year through questionnaires, interviews, and observations, and evidence generated naturally during the course of the project was also analysed. Each of the workshops at the end of year conference was observed, with attention given to indicators of the working relationship among the subgroup members. Two open response questionnaires were used. All the teachers present at the final full staff conference were invited to complete a questionnaire reflecting on their learning (41 returned—approximately half of those present, excluding the working group). Members of the working group completed a different questionnaire that, among other things, sought their views on the benefits and challenges of cross-phase collaboration (13 returned). Interviews were also conducted with five working group members, following up points in their questionnaires and exploring other aspects of the project.

Interviews were taped and transcribed, questionnaire data entered onto spread sheets, and observations recorded on a proforma. Data analysis was conducted by the three researchers/authors, first separately then together comparing interpretations.

## **12.4 Findings**

### ***12.4.1 Collaboration Generates Resources and Practices to Improve Learning***

The stimulus and opportunities offered by sharing practice in learning communities generated developments that ultimately impacted upon learning. Practices developed and resources created through collaboration among teachers were trialled and improved before being showcased at the cluster conference. The main developments are outlined in Table 12.2, together with the subject focus, numbers of teachers, school phase, and related principle(s). Each group had specific subject orientations and most generated more than one practical resource.

A key feature common to the work of groups A (subject focus Mathematics), B (subject focus English), and C (subject focus outdoor and practical learning) was the development of pupil self-assessment through their discussion of learning, achievements, and next steps in relation to clearly expressed success criteria. Group C and group E (English) both focused on peer assessment, either during the course of practical activities or, in the case of English, after a piece of writing had been drafted. Teachers in groups D (Science and Geography), along with the secondary colleagues in group A (Mathematics), concentrated on their response to pupils' work and developing ways for pupils to act on and learn from the feedback.

Teachers particularly appreciated innovations that had a clear impact on pupil learning and were adaptable across different age ranges and subjects. By collaborating and building on one another's ideas and practices, teachers found they did

not have to start from scratch and could trial new approaches with the confidence that they had others' support. Teachers' collective experiences with AfL enabled them to be mutual 'sounding boards' when discussing new ideas and to 'bounce ideas off each other' (Headteacher, primary school).

Through sharing teaching approaches and experiences, which led to key insights about AfL, project participants were able to improve their practice and draw the pupils themselves into the processes of assessment. For example, one primary school teacher said that the project had been tremendously helpful: as a classroom teacher he had improved the quality of his marking and feedback, and as a leader he ensured whole-school initiatives were clearly focused on learning (rather than performance). Another teacher reported how he had been having difficulty with trying to improve writing in his primary classroom, and it was through working with a group member from the secondary school that he had found a solution to the problem.

Project teachers were committed to using the practices and resources they had created and to developing the ideas even further. Feedback from colleagues who attended the conference was also very encouraging with many reporting that they were excited at the prospect of implementing new approaches in their classrooms.

#### ***12.4.2 Collaboration Aids Appreciation of Others' Contexts and Has an Affective Dimension***

Teachers admitted that prior to joining the project they had very little knowledge of one another's curricula and appreciated the opportunities the project afforded. They made links between what was occurring in quite different classroom settings, gained an understanding of how the same AfL strategy can be developed in both primary and secondary schools, and how teachers in both contexts can effectively support students' learning. For example, a secondary school teacher observed that he had developed a much clearer understanding of the prior experience pupils brought with them from primary school. These insights allowed teachers to plan their teaching more appropriately by focusing on pupils' prior experiences of assessments, easing school transfer for pupils, and aiding their learning.

The value and utility of cross-phase understanding was appreciated, especially as a means to achieving consistency in approaches as well as ensuring that learning and challenge were maintained. For many of the teachers of all phases and subjects, it was the insight into the similarities and differences of using AfL that was felt to be useful in giving a new understanding of approaches and areas of learning.

Primary teacher: Although we only met up four times over the year, I think those four times listening to each other talk were really good. The fact that all of the staff have come back to school and said things like 'I love this idea that I saw.' or that 'I really like what (someone) was saying about this.' Already that shows that it has improved understanding across phases.



One group observed each others' teaching across different settings to learn more about the different contexts, which was highly valued by all concerned. A primary school teacher commented that working cross-phase was 'a really good idea' as there was 'a sense of community amongst all schools.' The conference extended this opportunity, albeit to a more limited extent, to all the cluster's teachers. They learnt from colleagues, novice and experienced alike, who were teaching other age ranges and subjects.

Teachers appreciated the opportunity afforded by the working group to discuss and reflect on their assessment practices. A secondary school teacher shared that 'you do not have the opportunity to discuss your work with colleagues on a day-to-day basis because we (teachers) are all shut away in individual classrooms.' A sense of solidarity grew amongst teachers from different schools as they felt they were working in a collaborative, nonjudgemental way. For another teacher the project had helped develop a network of teacher colleagues, deepen understanding of principles underpinning assessment for learning, and reaffirm her moral purpose:

Primary teacher: You don't always have much time to see other people teach after your first year. I didn't really know what to expect from it, but what I've got out of it are the connections I've made with other people, reestablishing my understanding of why I'm doing what I'm doing.

Through sharing their AfL practices teachers felt affirmation for their own work and appreciation of the work of others. The collaboration increased empathy among teachers who reported feeling increased regard for colleagues in different subjects and age phases. Recognising the commonality of challenges faced—as exemplified by the comment 'there are similar issues between age groups but feedback is important throughout'—contributed to collegiality. Further evidence of the development of bridging social capital across teaching contexts came from a primary teacher who reflected on how useful it was 'to see someone who is a real expert in a subject area, but equally for the secondary school teachers [to see] we have so many different things in place to support so many different types of learners.' Secondary teachers were alerted to the depth of learning that takes place in primary schools, for example, in the quality and sophistication of children's writing.

Teachers also empathised and gave one another moral support when problems regarding AfL implementation (for example classroom time management, pressure of examinations) were shared. Being part of a learning community meant that teachers with less teaching experience could count on others and learn from their rich experiences of practising AfL. An early career primary teacher felt she 'really benefited from being in a group with really experienced teachers who appreciated the fact that it was my first year and I was a bit nervous and not as confident as the others. It was nice, because they really included me within the group.'

Project teachers who presented their work at the conference demonstrated a strong sense of teamwork and frequently referred to what they had achieved together. By the end of the year the groups were easy and familiar with one another, and teachers from different schools demonstrated an understanding of the disparate contexts in which their colleagues taught. Secondary school project teachers noted

how much they learnt about being student-focused from the early-years teachers in their group, while primary school teachers reflected on how subject-specific practices in secondary classrooms could also be useful for younger pupils.

At the conference there were lively discussions during various workshops as colleagues learnt about the variety of AfL practice across the cluster of schools. For example, a secondary teacher was particularly impressed by the productive use of peer assessment with very young children, while another reported that she was struck by ‘the difficulty of giving constructive feedback to KS2 students [pupils aged 7–11 years] for Maths.’ In one of the workshops project teachers were very open and spoke passionately about how much they had learnt from working group members who taught in schools across the primary–secondary divide. Colleagues who heard this were excited, curious, and in agreement that the ideas shared were useful across all age ranges. They were genuinely interested in how they could incorporate ideas into their own classroom assessment practices.

Teachers also reported that the conference helped them appreciate that implementing AfL is challenging regardless of school sector, and that there are considerable similarities in practice: ‘many teaching assessment strategies are cross-curricular and able to be used across the age ranges; primary and secondary have far more links than I realised.’ The way project teachers had worked together to develop ideas, and then collaborated in demonstrating their feasibility during the conference workshops, helped their colleagues see that AfL methods they may have thought suitable for only a specific age or subject have much more extensive applicability.

### ***12.4.3 AfL Principles Transcend Differences, Provide a Common Language and Promote Coherence***

For some teachers, the applicability of AfL strategies across subjects and ages was a revelation in itself. Differences in AfL practice were seen positively when connections were made to the three over-arching principles of promoting autonomy, making learning explicit, and focusing on learning rather than grades. These three principles transcend variations in practice among teachers of different subjects and ages. Innovations were seen to have common underpinnings, while unfamiliar practices were analysed for the principles they met, and then ‘struck a chord’ (secondary school teacher). The three AfL principles were cited by most teachers as key to their learning and development during the project, with individuals commenting on how the principles ‘really helped to clarify my thinking’ and ‘simplified thoughts’ helping to streamline ideas for the successful implementation of AfL.

According to the project teachers, the AfL principles provided a common language and promoted coherence within the group. They helped teachers gain an insight into the similarities and differences in teaching and learning across the key stages, highlighting, for example, that quality feedback was an issue for many. A primary school teacher shared how his group decided to focus on improving

writing quality: ‘All members in the group sat together around a table and had a “long long” conversation on what we had been doing to try to improve the quality of writing across the levels.’

The group then used their collective experiences to develop an effective feedback system for writing which was transferable across the age ranges. This happened across many groups where teachers found that they could reflect deeply about their own AfL practice based on the three key principles by learning from colleagues who taught other subjects and levels.

A primary school headteacher who was in one of the working groups asserted that ‘because staff vary in knowledge and understanding ... it’s really important that from the beginning, you set down the base of what the expectation is of AfL.’ The introduction of the three principles assisted teachers in focusing on what they sought to improve in their AfL practice. Many reported that they had made these principles key to all their lesson planning; comments included, ‘the focus is entirely on the learning not grades, marks or levels, and is all the better because of it!’ (secondary school teacher); and ‘in my role as deputy head [the principles have] also made staff meetings I have led and other whole school initiatives more meaningful and focused’ (Deputy headteacher, primary school). At the final conference, many of the other teachers commented positively on the principles, noting them as ‘key aspects I need to develop’ and saying they would be considering them in the future.

#### ***12.4.4 Multi-School Collaboration Presents Practical Challenges***

Even though the project was supported by all the schools in the cluster, there were considerable practical difficulties. The main challenge was the perennial one of time, manifested in different ways. For individuals, balancing time between teaching and other commitments and duties, whilst also trying to develop AfL practice and contribute to the project, was acknowledged as a particular challenge that was ‘part of life as a teacher.’ The two occasions when teachers from all the schools in the cluster met together took considerable planning and co-ordination, with everyone needing to commit to professional development activity at the same time in the same place. It is often hard to arrange a full staff meeting in a single school, much less for all the teachers from eight schools.

The four working group meetings required teachers from across the cluster to be able to meet together, necessitating close liaison and commitment. Similarly, enabling teachers to visit others’ classes for peer observation and additional sub-group meetings required coordination and multiple layers of arrangements. Additional expenses were incurred when cover had to be provided for teachers who were absent from their classrooms. Nevertheless, the teachers who were able to visit others’ classes considered it to be ‘100 times more useful than any other form of training or professional development’ (Deputy headteacher, primary school).

Teachers typically experience difficulties in finding the time and space to have professional conversations with colleagues in their own schools, let alone in other schools. Sometimes fleeting chances to communicate were grabbed, and partial solutions found, such as ‘conversations’ by email, but these had their own challenges and were seen by some as ‘convoluted’ when a ‘quick chat’ face-to-face would have been preferable.

Other pressures, such as those arising from external inspection and maternity leave, are not easily accommodated or necessarily foreseeable. Teachers from one school had to miss a group meeting since precedence had to be given to their own full staff meeting on an issue that had arisen, and with seven other schools involved it was not practical to rearrange the project meeting.

Nevertheless, despite the challenges high participation continued throughout the project. For at least one group member choice was a key factor in the success of the project, suggesting that teacher agency is crucial for making changes to practice. That agency includes the freedom to choose to participate and which practices to focus upon. For the teachers involved in the development project, it was this bottom-up choice that led to teams of teachers who were motivated towards making changes and working collaboratively.

Collaboration between such motivated teachers, all working to develop practice together, was highly valued and regarded as most beneficial to the extent that it was planned to continue with that model of CPD. The project’s success and teachers’ appreciation of its benefits generated the desire for more time to spend on this and similar professional development. The AfL project was only prioritised for one year, although the work linked closely to CPD foci in the following year when inter-school work concentrated on the implementation of a revised national curriculum and the associated guidance for ongoing assessment not to use grades. The secondary school also incorporated AfL into a new working group involving other teachers exploring how to help pupils develop a ‘growth mindset.’

## **12.5 Conclusion**

### ***12.5.1 Four Themes***

A number of conclusions may be drawn from the empirical data and can be grouped into four main themes:

- the benefits of cross-phase collaboration for developing AfL;
- the contribution of cross-phase professional development on AfL to pupil transfer;
- the value of principles to developing AfL practice in multiple settings;
- the role of leadership in supporting professional development across a cluster of schools.

Working group members were unanimous in their appreciation of the opportunity for cross-phase professional development, the reasons for which included, but went well beyond, ‘the positives of collaborative professional inquiry models’ (Kennedy 2014, p. 692) increasingly found in contemporary literature. Learning with and from each other, appreciating differences, sharing and developing practices, and finding ways of improving pupils’ learning were all highly valued by both primary and secondary teachers. This contrasts with Pedder and Opfer’s (2011) findings that suggested secondary teachers’ motivation for CPD participation is more likely to be career enhancement. The constituency of the working group resembled Lave and Wenger’s (1991) depiction of a community of practice in that it included novices (newly qualified teachers), experts (senior teachers and school leaders), and academics, with all parties learning with and from each other. Participants also developed reciprocal critical friendships (Swaffield 2004), with colleagues providing nonjudgemental feedback and acting as sounding boards.

Participants’ dialogue indicated that cross-phase CPD focused on AfL has particular benefits for easing pupil transfer from primary to secondary school. The working group as a whole and each of the subgroups resembled Mizelle’s (2005) ‘vertical team,’ through which teachers realised they faced very similar issues and were helped to refine and develop their AfL practices. They also came to understand more about pupils’ experiences and achievements beyond their own classes, as well as the expectations and demands in other lessons pupils attend before, after, and alongside their own. Activities focused on AfL are relevant to all teachers as AfL is cross-curricular, which may make organising cross-phase working easier than when it follows more traditional, subject-specific demarcations. Thinking about transfer to secondary school in terms of AfL also draws attention not just to the primary–secondary divide but also to the divisions among subjects in the senior school. It seems likely that if the practices of specialist subject teachers are more closely aligned, then a pupil’s move from a single class teacher in primary school to multiple teachers in secondary will be smoother. AfL acts as a *pedagogical unifier* across subjects and age ranges. A final way in which AfL may ease transfer is through its promotion of learning autonomy, which helps pupils to be more self-regulating and to see themselves as agents of their learning, able to make sense of apparently different learning environments.

Promoting learning autonomy, along with the other two principles of making learning explicit and focus on learning rather than grades, were the conceptual framework used throughout the project. Their introduction and promotion was perhaps the academic’s most significant contribution (beyond facilitating the whole process of collaborative inquiry), illustrating Lave and Wenger’s (1991) recognition of the value of different roles in a community of practice. The three principles encapsulating the essence of AfL provided a focus for the entire project, a framework for analysing what can appear very disparate practice (the ‘simplification’ mentioned by one teacher), and a common vocabulary. They were referred to throughout the meetings, were made explicit in all the end of year workshops, and proved to be a powerful scaffold for discourse (Swaffield 2006). Evidence from this small scale project strongly suggests that a framework of principles integrated into a

programme of professional learning and development considerably enhances its effectiveness. The three AfL principles were derived from the Learning How to Learn research (James et al. 2007), and as has been demonstrated are applicable across subjects, age ranges and contexts; in other words, they exhibit key characteristics identified by Bruner (1966)—power and economy. The formulation and use of principles to guide practice has a long pedigree in education: for example Stenhouse advocated their use in his seminal work in 1975, that included reference to Peters' (1959) discussion of principles of procedure. A more recent example is the Leadership for Learning framework with its five principles for practice developed through an international project (MacBeath and Dempster 2009) that have directly influenced policy and practice in many contexts across the world (Swaffield et al. 2014).

The project was led overall by a senior leader from the secondary school, in partnership with the head teachers of all the schools and a steering group. The senior leader liaised closely with the academic who had a major role in facilitating the working group and contributing to the joint staff meetings. Undoubtedly, the history and culture of collaborative working among the cluster schools and with the university helped the smooth running of the project. Another essential was the funding secured by the project coordinator which released teachers during the school day for working group meetings, provided refreshments for these and the joint cluster staff meetings, and enabled the academic's involvement. Limited funding only goes so far, and great benefit was derived from goodwill, resting on trust and commitment, and often made manifest by support in kind (for example subgroups' additional meetings, rooms, and research). Leadership decisions to include teachers across the spectrum of experience and seniority both built on and enhanced trust, and saw colleagues ranging from a newly qualified teacher to a head teacher all learning from and with each other. This developed all of Putnam's (2000) and Woolcock's (2001) forms of social capital—bonding, bridging, and linking.

### ***12.5.2 The Challenges for Implementation***

Notwithstanding all the advantages of cross-phase AfL-focused professional development, there are considerable challenges.

- *Collaborative time.* It takes time to develop the understanding and trust needed for collaborative professional learning, especially with colleagues from different contexts.
- *Duration.* One academic year is too short to embed changes in AfL practice, yet other topics (albeit related) arise deserving of attention.
- *Funding.* Enabling teachers to meet together, especially when they come from different schools, costs money and limits the number who can be directly involved.

- *Coordination and goodwill.* Overall coordination is best done by one or two people, and goodwill from everyone helps overcome the inevitable difficulties.
- *Unifying principles* are needed to bring conceptual clarity and coherence, and to aid dialogue of what could otherwise be seen as widely disparate and unrelated practices.
- *Sustaining and scaling* remain perennial challenges.

### **12.5.3 Recommendations**

The research reported here involved just one secondary school and its feeder primary schools, so considerable caution is necessary when making generalised recommendations. Moreover, these schools had a history of collaboration, school leaders valued and supported the project, and the local culture was conducive to cross-school working. Nevertheless, the findings of this study suggest a number of points for consideration by teachers, policymakers, and particularly school leaders more widely.

Since it appears that teachers' AfL practices can be enhanced through working with colleagues in different contexts, teachers generally might be advised to take and create every opportunity to do so. This is much easier when such working arrangements are valued and facilitated by school leaders, ensuring practical coordination and providing necessary funding.

Research-generated, parsimonious principles may well enhance the efficacy of professional learning activities. The AfL principles identified by James et al. (2007) certainly resonated with the teachers involved in this project and thus gained further endorsement. It seems likely therefore that other teachers will also find that the three principles of making learning explicit, promoting learning autonomy, and focusing on learning rather than grades, provide a powerful conceptual framework. For school leaders these principles offer a warranted structure for the planning, review, and evaluation of AfL practices and policy. Senior teachers who were members of the working groups attested to the value of the principles, suggesting that school leaders would be advised to practise principle-directed AfL themselves, thus deepening their appreciation and understanding of assessment for learning and its underpinning principles.

The potential of AfL as a pedagogical unifier assisting pupil transfer between schools, and the benefits of teachers working with colleagues from other schools, indicate that cross-phase AfL-focused professional development is a commendable practice. It is predominantly school leaders who are in the position to set up local working groups, recognising that it takes considerable time to establish the necessary trust and understanding.

Whilst much can be achieved at the local level, national policy and guidance influence the prevailing culture. Policymakers could encourage ongoing, cross-phase collaborative development work focused on assessment for learning,

not least by avoiding frequent changes in national assessment arrangements so that school leaders and teachers can concentrate on AfL. Given the encouraging outcomes of this small scale study, further research into similar cross-phase professional development would be welcome.

## References

- Beaumont, C., Moscrop, C., & Canning, S. (2014). Easing the transition from school to HE: Scaffolding the development of self-regulated learning through a dialogic approach to feedback. *Journal of Further and Higher Education*, Published online October 09, 2014. doi:10.1080/0309877X.2014.953460
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, 33(8), 3–15.
- Bruner, J. (1966). *Towards a theory of instruction*. Cambridge, Massachusetts: Harvard University Press.
- Clausen, K. W., Aquino, A.-M., & Wideman, R. (2009). Bridging the real and ideal: A comparison between learning community characteristics and a school-based case study. *Teaching and Teacher Education*, 25(3), 444–452.
- Cooper, C., & Boyd, J. (1998). Creating sustained professional growth through collaborative reflection. In C. M. Brody & N. Davidson (Eds.), *Professional development for cooperative learning: Issues and approaches*. Albany: State University of New York Press.
- Craven, R. G., Marsh, H. W., & Debus, R. L. (1991). Effects of internally focused feedback and attributional feedback on enhancement of academic self-concept. *Journal of Educational Psychology*, 83(1), 17–27.
- Darling-Hammond, L. (1996). What matters most: A competent teacher for every child. *The Phi Delta Kappan*, 78(3), 193–200.
- Earl, L. (2012). *Assessment as learning* (2nd ed.). Thousand Oaks, California: Corwin Press.
- Galton, M., Gray, J., & Rudduck, J. (2003). *Transfer and transitions in the middle years of schooling (7–14): Continuities and discontinuities in learning* (No. 443). Annesley: Department of Education and Science.
- Gardner, J. (2006). Assessment for learning: A compelling conceptualization. In J. Gardner (Ed.), *Assessment and learning*. London: Sage.
- Gardner, J. (2010). Teachers as self-agents of change. In J. Gardner, W. Harlen, L. Hayward, G. Stobart, & M. Montgomery (Eds.), *Developing teacher assessment*. Maidenhead: Open University Press, McGraw-Hill Education.
- Ginsburg-Block, M. D., Rohrbeck, C. A., & Fantuzzo, J. W. (2006). A meta-analytic review of social, self-concept, and behavioral outcomes of peer-assisted learning. *Journal of Educational Psychology*, 98(4), 732–749.
- Hargreaves, E. (2005). Assessment for learning? Thinking outside the (black) box. *Cambridge Journal of Education*, 35(2), 213–224.
- James, M., Black, P., Carmichael, P., Drummond, M.-J., Fox, A., MacBeath, J., et al. (2007). *Improving learning how to learn: Classrooms, schools and networks*. London: Routledge.
- James, M., & McCormick, R. (2009). Teachers learning how to learn. *Teaching and Teacher Education*, 25(7), 973–982.
- James, M., & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *Curriculum Journal*, 17(2), 109–138.
- Katz, S., & Earl, L. (2010). Learning about networked learning communities. *School Effectiveness and School Improvement*, 21(1), 27–51.



- Kellard, K., Costello, M., Godfrey, D., Griffiths, E., & Rees, C. (2008). *Evaluation of the developing thinking and assessment for learning development programme*. Birmingham: Report for BMG research.
- Kennedy, A. (2014). Understanding continuing professional development: The need for theory to impact on policy and practice. *Professional Development in Education*, 40(5), 688–697.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation* (1st ed.). Cambridge, UK: Cambridge University Press.
- Little, J. (2005). Professional learning and school-network ties: Prospects for school Improvement. *Journal of Educational Change*, 6(3), 277–284.
- MacBeath, J., & Dempster, N. (Eds.). (2009). *Connecting leadership and learning: Principles for practice*. Abingdon, UK: Routledge.
- Miller, D., & Lavin, F. (2007). ‘But now I feel I want to give it a try’: Formative assessment, self-esteem and a sense of competence. *The Curriculum Journal*, 18(1), 3–25.
- Mizelle, N. B. (2005). Moving out of middle school. *Educational Leadership*, 62(7), 56–60.
- Pedder, D., & Opfer, V. D. (2011). Are we realising the full potential of teachers’ professional learning in schools in England? Policy issues and recommendations from a national study. *Professional Development in Education*, 37(5), 741–758.
- Peters, R. S. (1959). *Authority, responsibility and education*. London: Allen & Unwin.
- Putnam, R. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster Ltd.
- Rosenholtz, S. (1989). *Teacher’s workplace: The social organization of schools*. New York: Longman.
- Sato, M., Coffey, J., & Moorthy, S. (2005). Two teachers making assessment for learning their own. *Curriculum Journal*, 16(2), 177–191.
- Smith, K. (2011). Professional development of teachers—A prerequisite for AfL to be successfully implemented in the classroom. *Studies in Educational Evaluation*, 37(1), 55–61.
- Smith, K., & Engelsen, K. S. (2013). Developing an assessment for learning (AfL) culture in school: The voice of the principals. *International Journal of Leadership in Education*, 16(1), 106–125.
- Stenhouse, L. (1975). *An introduction to curriculum research and design*. London: Heinemann.
- Suurttamm, C., Koch, M., & Arden, A. (2010). Teachers’ assessment practices in mathematics: Classrooms in the context of reform. *Assessment in Education: Principles, Policy & Practice*, 17(4), 399–417.
- Swaffield, S. (2004). Critical friends: Supporting leadership, improving learning. *Improving Schools*, 7(3), 267–278.
- Swaffield, S. (2006). Scaffolding discourse in multi-national collaborative enquiry: The carpe vitam leadership for learning project. *Leading & Managing*, 12(2), 10–18.
- Swaffield, S. (2011). Getting to the heart of authentic assessment for learning. *Assessment in Education: Principles, Policy and Practice*, 18(4), 433–449.
- Swaffield, S., Dempster, N., Frost, D., & MacBeath, J. (Eds.) (2014). *LfL travels*. InForm 17. University of Cambridge.
- Swaffield, S., & MacBeath, J. (2006). Embedding learning how to learn in school policy: The challenge for leadership. *Research Papers in Education*, 21(2), 201–215.
- Tierney, R. D. (2006). Changing practices: Influences on classroom assessment. *Assessment in Education: Principles, Policy & Practice*, 13(3), 239–264.
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge, U.K.: Cambridge University Press.
- William, D. (2007a). Changing classroom practice. *Educational Leadership*, 65(4), 36–42.
- William, D. (2007b). Content then process: Teacher learning communities in the service of formative assessment. In D. B. Reeves (Ed.), *Ahead of the curve: The power of assessment to transform teaching and learning* (pp. 183–204). Bloomington, Indiana: Solution Tree.
- William, D. (2009). *Assessment for learning: Why, what and how? An inaugural lecture by Dylan William*. London: Institute of Education.

- Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *Australian Educational Researcher*, 40(2), 241–256.
- Woolcock, M. (2001). The place of social capital in understanding social and economic outcomes. *ISUMA: Canadian Journal of Policy Research*, 2(1), 11–17.

## Chapter 13

# Implementing High Quality Assessment for Learning: Mapping as a Professional Development Tool for Understanding the *What to Learn, Why to Learn It, and How to Learn It*

Maria Araceli Ruiz-Primo

**Abstract** Effective implementation of assessment for learning requires that teachers fully grasp the ‘big picture’ for the unit and lessons/modules that they will be teaching their students over a period of time, including all the related activities. This means understanding how various blocks of knowledge fit together, especially at the higher levels, which students will use to build conceptual frameworks focused on principled understandings—the big ideas. This chapter recognizes the importance of understanding the *what*, the *why*, and the *how* of what teachers need to teach and student to learn. It offers a proven strategy designed to help teachers develop the big picture through an interactive process called *mapping*. Different empirical studies have demonstrated that the mapping process can lead teachers to discover new and complex insights about *what* they are to teach and *why* this content is important. Further, the mapping process helps them become familiar with the kinds of events and phenomena necessary to understanding *how* to ensure that their students achieve the expected learning outcomes.

### 13.1 Introduction

Critical to sound assessment for learning is the ability of teachers to implement strategies that will allow them to effectively: (1) share the learning goals, expectations, and criteria with their students, (2) elicit information critical to determining where students are in their learning, (3) analyze and meaningfully interpret the collected information to make it usable for them and their students, and (4) act upon or use the collected information in ways that will directly support the students’ learning. While these activities generally tend to be viewed as easy to implement,

---

M.A. Ruiz-Primo (✉)  
Stanford University, Stanford, USA  
e-mail: aruiz@stanford.edu

research shows that teachers often struggle simply to identify the learning goals of what they are teaching (Dwyer 1994; Ruiz-Primo and Li 2004; Ruiz-Primo et al. 2010). And even when teachers can articulate the relevant learning goals, they commonly have difficulty identifying, selecting, or developing meaningful instructional activities that are cognitively demanding and properly sequenced for achieving those goals (Lingard 2007) or explaining why an instructional activity is important to implement. These limitations are the result of poorly written curricula, time constraints, and ambiguous state and national standards.

The purpose of this chapter is to present and describe a process, called *mapping* (Ruiz-Primo and Li 2009), previously proven to enhance teachers' understanding of learning goals and how to achieve them. Both the nature and impact of this process are relevant to Part II of this book. First, for mapping to be successful, it needs to take place in a collaborative social context in which teachers have conversations about *what* learning needs to happen in a study unit, *why*, and *how*. Together they construct such understanding by moving back and forth among the lessons to be taught. Second, although the mapping process was an activity in a research grant funded by National Science Foundation for developing instructionally sensitive assessments (Ruiz-Primo and Li 2008), the participating teachers treated it as a professional development experience, even though it was never designed or presented to them as such. The mapping process helped acquaint the teachers with the essential tools, physical and cognitive, that would allow them to analyze more critically the content and the sequence of what they had to teach. They were better prepared to identify gaps and to make informed decisions about how to fill and resolve these issues. Mapping is proposed as a strategy to improve assessment competencies that teachers need to acquire in order to better implement assessment for learning with their students. If teachers do not have clarity about the learning goals being pursued the quality and impact of formative assessment on students' learning is jeopardized. It is a pre-requisite to engage students in productive tasks and dialogues.

The chapter is organized in five sections. Section 13.2 explains why it is important for teachers to better understand what is to be taught. Section 13.3 describes *mapping*, a process that can help teachers to develop this increased understanding. Section 13.4 suggests some effects on the implementation of assessment for learning when teachers fully understand what is to be taught. Section 13.5 closes the chapter with some conclusions.

## **13.2 The Importance of Understanding the 'What', the 'Why', and the 'How'**

A critical prerequisite for implementing assessment for learning is sufficient clarity about the learning goals to be achieved. Teachers must know where they want to take their students—that is, what students need to learn. Only with clear learning

goals can teachers determine what information they need to collect, the most appropriate strategies for gathering the information, what evidence will show that learning has taken place, what they need to pay attention to, and when gathering information should be a formal process. More specifically, teachers should clearly understand:

- *What* is to be learned (i.e., the learning goals that need to be achieved by the end of a unit/module/topic).
- *Why* this learning is important for students in the context of the unit/module/topic or curriculum, and even beyond.
- *How* these learning goals are to be achieved.
- *What* is important to pay attention to (notice) in students' responses to understand where they are in their understanding.
- *How* to know that the desired learning has actually been achieved.

Evidence that teachers have acquired a sufficiently deep understanding of the 'whats, whys, and hows' should be reflected in their ability to respond to questions such as, Why am I teaching this content? Why is it important? Why are the activities sequenced in the way that they are? How does each activity contribute to achievement of the overall learning goal(s)? What critical foundational blocks of knowledge, practices, or skills need to be established to achieve the learning goals? What specific evidence will show that those foundational elements have been built? What evidence is needed to demonstrate that the learning goals have been met?

### ***13.2.1 What Learning Is to Be Achieved?***

My colleagues and I (e.g., Ruiz-Primo and Li 2009; Ruiz-Primo et al. 2012) contend that deep understanding of learning goals encompasses more than the list of learning objectives or learning targets that typically appear at the beginning of a unit, module, or chapter. Meaningful learning goals will reflect *big ideas*, complex ideas that can challenge students' thinking and significantly extend their capabilities and understandings (Eraut 1997). Big ideas reflect the principled understandings that typically form the foundation for more advanced learning (Brown et al. 1986; Windschitl n.d.). They are characterized by higher explanatory power and flexible application (Brown et al. 1986). This kind of learning should make a difference in how learners think and perform after having learned something new. If a learning goal does not change or alter thinking or reasoning, then there is no true learning (Eraut 1997). For example, will knowing simply the definition of density alter students' thinking about why things sink or float? Our research provides evidence that it is not the case (Ruiz-Primo et al. 2010).

When students demonstrate knowledge about something, it is presumed that they have constructed mental structures (principled understanding) that allow them to deploy the relevant knowledge demanded by a new task (Hickey and Pellegrino 2005).

This is not possible when knowledge structures are based mainly on narrow solutions or isolated rules (bits and pieces of information) that obstruct rather than facilitate transfer of learning (Brown et al. 1986). Instead, by deliberately aligning instruction to selected big ideas, it is possible to connect concepts that originally seemed disconnected. This may also lead to highlighting actual occurrences of the new idea(s) or explaining their extension through interesting, new reasoning. It is not by accident that high-quality national and state standards reflect big ideas rather than individual concepts.

Big overarching ideas help to organize and focus the learning goals that teachers need to achieve with their students. As expressed by Windschitl (n.d.), these ideas are constructed; they emerge as a result of careful thinking over a period of time. Arriving at big ideas requires deliberate examination and consideration of the links between relevant concepts, processes, and skills. They require going back and forth among the different elements of content that need to be taught until the connections among the elements are fully understood.

### ***13.2.2 Why Is This Learning Important?***

Understanding what to teach is still not enough. Teachers must also be clear about why it is important for their students to learn this content. Armed with this additional understanding teachers can more successfully share with and clarify for students the intended learning and its importance. They can ensure that their students will see connections that might otherwise be overlooked. Students need these connections, in turn, to build new connections (frameworks) that will make their knowledge more broadly applicable. In short, students will better understand why they are learning something, why they are doing what they are doing, and how this work/activity helps them to become successful learners.

### ***13.2.3 How Should the Learning Goals Be Achieved?***

Understanding how teachers and students are to achieve the given learning goal(s) requires tracking individual lessons/activities within each unit/module/chapter. This involves identifying: (1) what is to be learned for each lesson, (2) which instructional activities (experiences) are critical to helping students achieve the targeted learning, (3) the optimal sequence in which these activities should be implemented, and (4) what to look for during each activity that indicates whether students are learning as expected. Being aware of how each activity contributes to achieving the learning goals enables teachers to help students understand why it is worth doing something and what they need to learn from what they do: ‘everything included in a curriculum should be included because it is worth learning for reasons that can be understood by the learners...’ (Brophy 1999, p. 80).

Identifying how students' learning is being constructed based on the instructional activities they are experiencing helps to identify a learning path—that is, how learning progresses over the course of the instruction as guided by the sequence of activities and lessons. Ultimately, familiarity with the learning path will make it possible to identify (1) the instructional experiences critical to achieving the learning goals for the individual lessons and the overall unit/module, as well as (2) the critical junctures at which daily practices for collecting information need to be augmented by more formal strategies.

It is essential that instructional activities and tasks are rooted in worthwhile learning and purposes. Teachers (and students) should be clear about how each activity contributes to achieving the learning goal. We have observed teachers implementing mechanically the activities suggested in science modules that trivialize the original purpose because they do not have a firm sense of what students need to learn. For example, rather than focusing on whether students understand why it is important to keep a variable constant and to manipulate another one, teachers focus on making sure the students simply draw the system (e.g., flippers) correctly with appropriate labels. Thus, for each activity it is important to consider two questions: What is to be learned (or practiced) in this activity? In what respect(s) does this learning contribute to achieving the larger learning goal(s)?

### ***13.2.4 How to Know that the Desired Learning Has Been Achieved?***

Knowing what to look for as evidence that learning is occurring helps teachers to be purposeful in selecting, modifying, adapting, or designing activities and to ask questions that can reveal students' thinking relative to the desired learning. For example, mathematics problems should be designed to make it possible for teachers to easily recognize the different ways in which students can solve problems. Well-designed classroom tasks will readily reflect the struggles students are having, information that teachers can use to help their students. High quality tasks can help teachers to judge the quality of students' thinking from the work they are doing or their responses. We have observed teachers who know exactly what to look for while students are working (individually, in pairs, or small groups). They gather information while walking around students' tables or desks, and they immediately act upon that information at the individual level. We have learned that teachers with expertise in implementing assessment for learning tend to be naturally adept at reaching many individual students with timely feedback (e.g., with a finger pointing out at critical steps of the procedure, asking a critical question) during everyday class activities.

### 13.3 Mapping: A Process for Coming to Understand the What, the Why, and the How

Knowledge of a curriculum requires more than familiarity with the sequence of instructional activities. It involves distinguishing between what is and is not critical to students' learning. It involves understanding why something comes first and something else comes later. Curriculum understanding further involves specifying situations in which students are learning the concepts and knowing how such situations can be modified to assess students' ability to apply what they have learned to new problems in novel contexts.

Needless to say, teachers make professional judgments and adjust the curriculum to their particular context and students. The issue is whether such adjustments are made with full understanding of the implications of those adjustments. Without such understanding the adjustments remain superficial. In one of our projects (Ruiz-Primo and Li 2008), when we asked teachers why they did not implement certain activities (usually the final lesson) in a unit, they simply cited 'time constraints', without referring to the relative importance of that lesson to achievement of the learning goal.

Mapping nudges teachers to continually ask themselves, What needs to be taught and why, Why does Activity A come before Activity B, What evidence do I need to know that my students have achieved the learning target of this activity? With this kind of complex, layered understanding, teachers are ready to guide their students to ask themselves, What needs to be learned? What evidence do I need to show that I have accomplished this learning?

#### 13.3.1 Mapping as an Iterative Process

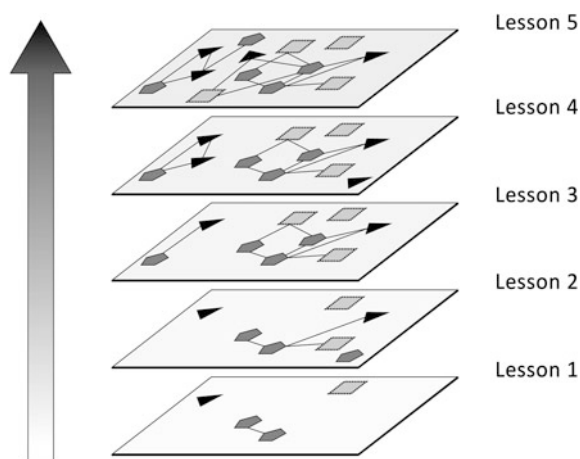
All curricula are organized in some way but not the same way. For example, science curricula in the United States are usually organized around units or modules that involve investigations. Within an academic year more than one unit or module is implemented. (For simplicity purposes, hereafter I use the term 'unit' to refer to modules or chapters or topics, and 'lesson' to refer to the smaller parts of a unit.) Mapping the 'curriculum' is carried out by *mapping the units* that make up a curriculum. Both the unit and the lesson levels are considered in what follows.

*Mapping* is an *iterative process* that allows teachers (the mappers) to move back and forth between two levels of analysis—the lesson and the unit as a whole—to ultimately discern the essence of a unit. Mapping helps teachers to grasp at a deeper level the essence of the *intended curriculum* by identifying the *concepts, processes, problem-solving approaches, proof schemes, or principles* critical across each lesson of the unit and how they become interrelated from one lesson to the next.



Understanding how teachers and students are to achieve the unit learning goal(s) requires *tracking the individual lessons within each unit*. For each lesson, the learning pursued (i.e., the ‘learning target’ at the lesson level) should be identified together with the instructional activities that need to happen for that learning target to be achieved. As the teachers (mappers) analyze each lesson, they need to go back and forth among lessons to review and make any necessary changes to the learning targets and the critical activities. The reason for these reviews is simple: the role that something plays in future lessons is understood better as the analysis across lessons moves forward. What initially was considered important while analyzing Lesson 1 may appear unimportant once the teachers begin analyzing subsequent lessons. Or, on the contrary, what seemed unimportant at the beginning might assume critical importance in a future lesson. Analysis at a higher level gives teachers a fuller, more accurate perspective on what does and does not matter relative to the unit learning targets. This back and forth analysis and adjustment makes mapping an *iterative* process. The more fully the teacher understands a lesson and how it contributes to achieving the unit learning goal, the more closely in tune will be the learning targets and critical activities across lessons.

Mapping assists teachers to identify: (1) The essential ideas, or learning goals, around which the unit appears to be constructed. For example, Fig. 1 shows two concept networks—conceptual frameworks—by Lesson 5 which should be the focus of the instruction. And (2) the concepts that, in the end, are not clearly related to other concepts and therefore should not be considered fundamental (i.e., they are peripheral or secondary). These concepts might be introduced in a lesson, but they may not be relevant or connected across other lessons (see the pentagon in the lower right corner of Lesson 2 and the triangle on Lesson 4).



**Fig. 13.1** Graphical representation of the interrelations among unit elements (e.g., concepts or processes) across lessons within a unit

Alternating reviews of the lessons help to identify both *what* is critical, *why* that learning is critical, and *what* evidence is needed to show that the learning occurs. The mapping process clarifies for teachers the building blocks essential to future learning—that is, the most critical knowledge and skills that the students need to acquire as they complete a unit.

### 13.3.2 Mapping Components

In our assessment research project (Ruiz-Primo and Li 2009) teachers were asked during mapping to keep in mind the critical question, ‘What is the lesson asking the teacher and the students to focus on and to learn from the perspectives of both scientific knowledge and scientific practices?’

To help the teachers respond to this question, the mapping process we proposed (Ruiz-Primo and Li 2009) included three components: (1) a framework for thinking about the science units—what should be captured and how, (2) a unit map—an artifact (tool) that helped mappers (teachers) to track their thinking around each lesson, and (3) a collaborative context that stimulated discussions about the science module—constructing conversations among experienced teachers who had taught the same science module.

The *unit map* helped to capture the group discussions and to synthesize the teachers’ experiences, knowledge, practices, and beliefs. The unit map is a matrix in which rows are the lessons and columns are the unit elements to be captured.

The framework originally involved tracking eight elements of each lesson within a unit (Ruiz-Primo and Li 2009): (1) learning targets for the lesson in terms of scientific knowledge and scientific processes; (2) cognitive demands that teachers and students needed to engage with during the lesson; (3) activities critical to achieving the lesson’s learning targets; (4) documentation (i.e., students’ products) required from students; (5) materials used; (6) graphical representations; (7) vocabulary used in this lesson/activity/investigation; and (8) any concerns teachers may have had about misconceptions shaped by students’ experiences. It is important to note that all these elements were important for the assessment project in which the mapping was embedded. However, a simplified mapping process can focus on fewer elements and still achieve its purpose. Table 13.1 provides an example of an adaptation of the unit map used in our assessment project.

The mapping process requires a social context made up of small groups of teachers, three to four in each group (Ruiz-Primo and Li 2009), who treat the unit map as a tool to focus their discussions in a productive manner.

Mapping can be a rich shared learning experience. During the mapping process participants contribute their own experiences of teaching the unit as well as their conceptual understanding of the unit. For each column of the unit map, teachers are asked to record what should be included from the unit or to generate an interpretation beyond what was specified in the unit.

**Table 13.1** Unit map (adapted from Ruiz-Primo and Li 2009)

Lesson	Learning Targets		Thinking Involved				Critical activities	Documentation
	Scientific knowledge	Scientific practices	Describing	Building explanations	Making connections	...		
1								
2								
...								
<i>n</i>								

Defining the learning targets guides the discussion of the unit lessons. Teachers are asked, however, not to define the learning goals of the unit (the big ideas) until all of the lessons have been analyzed and the learning targets of the lessons identified.

The learning targets for each lesson drive the discussion in two ways (Ruiz-Primo and Li 2009). First, they define the activities critical to reaching the targets at the end of the lesson. And, second, they make it possible to weigh the value of the various activities (i.e., what is to be learned in this activity that contributes to achievement of the learning target for this lesson?). As mentioned, the learning targets are iteratively revised as the teachers improve their understanding of the intended learning targets. This understanding emerges as the result of the discussions of the interplay among the other elements of the matrix and based on experiences from the activities.

The tangible outcome of the process is a *completed unit map* that reflects the negotiations among teachers (mappers) about the learning goals to be achieved. The map also indicates the critical areas of focus, the reasons for this focus, and the kinds of evidence about the students' learning that should be considered. One teacher mentioned to us that he used the completed unit map as his planning tool.

### 13.3.3 How to Map

Over a period of one to two days, teachers move through the process in the following manner:

1. Collective reading of each lesson of the unit, tentatively identifying the learning target based on their experiences teaching the unit.
2. As they reread the unit, they are asked to consider both their own and the students' perspectives.
3. When they teach the lessons, teachers are asked how the learning targets for each lesson reflect a trajectory of students' learning progress within the unit being mapped. They should recognize how the learning targets at the lesson level are part of a bigger story in the development of students' learning within the unit and, ultimately, across units. Being aware of the learning trajectory

assists teachers in making informed decisions about when to use a formal assessment at critical junctures during the unit (Ruiz-Primo et al. 2010).

4. Once all lessons are mapped and the learning targets by lesson are finalized, teachers collectively determine the learning goal of the unit.
5. In optimal situations, more than one group of teachers maps the same unit. If this is the case, the final step in the mapping process is to compare the groups' unit maps. Additional negotiation defines a final map to reflect the essential learning goal(s) of the unit and the learning targets by lesson, as well the major activities involving teachers and students to achieve such learning goals. Critical to the comparison is to define the unit learning goals as big ideas. Table 13.2 shows the difference between learning objectives proposed in some of the mapped science units and the big ideas that emerged from the mapping process.

### ***13.3.4 Impact of the Mapping Process on Teachers' Understanding of the Intended Curriculum: Examples of Empirical Evidence***

Curriculum mapping has been shown to help teachers improve their understanding of the targeted/desired learning goals, why these goals are important, how they can be achieved, and what evidence should be considered for determining whether students actually achieve these learning goals.

The mapping process and examples described here focus on our experiences implementing the mapping process on three occasions in the context of a science assessment project (Ruiz-Primo and Li 2008). I believe, however, that the process can be readily adapted to other disciplines. The mapping process developed by Ruiz-Primo and Li (2009) was implemented with 34 elementary school teachers in five school districts in the U.S. Midwest. These volunteer teacher participants were expected to teach the same science units during the school year. The process was replicated with four different science units over three years.

For three of the units we conducted the mapping process with two groups per science unit. Analysis of videotapes collected for the mapping of two units showed how the mapping process influenced teachers' understanding of the learning goals at the unit/module level (Giamellaro et al. 2011a).

Analysis of transcriptions of the videotapes conducted with two coders (Cohen's Kappa = 0.83) revealed that the mapping process helped teachers to identify, review, and revise the learning targets and the learning goals of the mapped units. Teachers often quickly accepted a lesson's learning target that, following analysis of later lessons, led them to realize that the initial learning target was not actually most appropriate for a given lesson. The number of iterations and the number of the curricular elements (columns in the unit map) discussed were considered indicators of the difficulty the teachers encountered when (1) identifying and articulating the learning targets and the learning goals and (2) explaining how the activities

contributed to achieving the pursued learning. While some lessons required only one iteration, others required up to five. Giamellaro et al. (2011a) found in the mapping sessions analyzed that teachers initially appeared to identify a high number of learning targets for a science unit (e.g., 18 learning targets across five lessons), but this number dropped significantly as the mapping proceeds (e.g., from 18 learning targets, teachers decided to adopt only 6 that they found to be critical for achieving the unit learning goals).

Another positive outcome was that teachers became aware of the trajectory of the students' learning and could see how the learning target at the lesson level led to achieving the unit learning goal. 'Understanding the end goal, the teachers were better able to vet the importance of certain curriculum parts' (Giamellaro et al. 2011a, p. 10). For example, teachers realized that some in their group never paid attention to the last lesson of a module, which indeed was essential to achieving the learning goals. The teachers also were critical of how certain learning goals proposed in the unit could not be achieved with the actual opportunities students received during the unit (see Giamellaro et al. 2011a). By examining the characteristics of the instructional activities through different lenses the teachers were able to distinguish between activities that truly contributed to students' learning and those that were more 'for students to have fun'.

In another study (Giamellaro et al. 2011b), we showed how the mapping process could reveal teachers' misconceptions regarding the content of the science units. We tracked seven fifth-grade teachers using videos recorded during the mapping sessions. Analysis of the videos showed that the mapping process exposed numerous teacher misconceptions, some of which were addressed by the group, while others required a content expert to join the group and lead a more in-depth discussion.

Finally, we have evidence that as a result of their participation in this mapping process some of the teachers implemented improved practices of assessment for learning, as shown by the coding of the videos taken during implementation of the mapped units (Ruiz-Primo et al. 2014). The link between the assessment for learning practices with students' performance is not perfect in every respect, but it is still encouraging (Lan et al. 2012; Wang et al. 2012).

### **13.4 Implications of Better Understanding What Is Taught**

As mentioned above, assessment for learning cannot be properly implemented without clarity about the learning goals being pursued and the instructional activities critical to achieving those learning goals. The mapping process has been proposed as an effective tool to develop a fuller and more accurate understanding about the what, the why, and the how. By acquiring this deeper understanding

**Table 13.2** Difference between learning objectives and big ideas. Example of a science unit on ‘Environments’

Learning objectives	Big ideas
Purposes <ul style="list-style-type: none"> <li>• Observe and compare organisms</li> <li>• Identify factors that make up a terrestrial environment</li> </ul> Scientific concepts <ul style="list-style-type: none"> <li>• Everything that surrounds organisms makes up the organism’s environment</li> <li>• An environment factor is one part of the environment. It can be living or nonliving</li> </ul>	<ul style="list-style-type: none"> <li>• For any particular environment, some kinds of plants and animals (living organisms) thrive (flourish), some do not live as well, and some do not survive</li> <li>• Changes (in environmental factors) in the organism’s habitat (environment) are sometimes beneficial to it and sometimes harmful</li> <li>• For any organism, there is a set of conditions (environmental factors) that are optimal for that organism’s survival, growth and reproduction</li> </ul>

teachers have the potential to develop high quality formative assessment strategies such as the ones mentioned below.

### ***13.4.1 Clarify the Learning Goals at Different Curricular Levels: Unit, Lesson, and Instructional Activity***

The degree of specificity of learning goals will vary depending on the curricular level: unit, daily lesson, instructional activity. For example, a day’s lesson and instructional activity require more specific learning goals than goals conceptualized as big ideas. How learning goals are articulated further depends on the *purpose* of an instructional activity. Defining such purpose requires the teacher to ask herself several questions: What do I expect students to learn by the end of this activity? How can I convey these goals to my students and also share them with others? The mapping process can help teachers to find answers to such questions.

Spending even a minute to communicate the purpose of an activity or how a day’s activities can contribute to progress in a learning path can ensure that students know where they are heading and assists them in making connections that otherwise might be missed. Such communication is a critical factor in leading students to attend to what they *need to learn* from an activity rather than just focusing on what they *need to complete*. In other words, students will learn more effectively if they fully understand what, why, and how they need to learn, as well as the kind of evidence that will demonstrate successful learning.

Learning goals should reflect the student learning trajectory at least within a unit but preferably also across units. Learning goals at the instructional activity level (task purpose) should be recognized as part of the larger picture in the development of students’ learning. Connecting learning with previous learning goals or activities can remain superficial (e.g., Do you remember when you talked about ‘x’ last year

in your science class?). But, making these connections can become complex and challenging if the stages in the trajectory of students' learning are purposefully linked.

### ***13.4.2 Design Activities that Support Students in Making Their Thinking Explicit***

Instructional activities can range from identifying critical questions to adapting the unit's instructional activities in ways that more accurately reflect what students are thinking and their level of understanding. Critical questions that push students' thinking yield the most information about students' thinking and are more likely to lead to more concrete actions from teachers to improve students' learning (Ruiz-Primo 2011; Ruiz-Primo and Furtak 2006, 2007). Most units, however, at least in science education, include questions that actually have proved to be irrelevant to achieving the unit learning goal (Ruiz-Primo et al. 2013). It is therefore left to the teachers to formulate relevant questions that actually support students in making their thinking explicit and which can lead to discussions that advance students' learning. Identifying such critical questions in the unit is possible only with clear unit learning goals.

Important to keep in mind is that increasing the percentage of assessment questions that require more than simple recall is an effective strategy for gathering information, because it nudges students to make their thinking explicit and guides the teacher toward appropriate strategies for helping those students. We have found that all teachers ask factual questions, but, unfortunately, not all teachers ask questions that require students to explain their ideas in explicit terms, their thinking, or how they arrived at their decisions (Ruiz-Primo and Furtak 2006, 2007; Ruiz-Primo et al. 2014). Teachers thus need to be aware that increased interactions that are true dialogues have been shown to improve students' learning and test performance (Mercer et al. 2004; Ruiz-Primo and Furtak 2006, 2007; Wiliam 2011). 'Good diagnosis relies on rich questions that elicit learners' higher-order thinking' (Stobart 2014, p. 118).

Similarly, directing a critical eye toward the instructional activities proposed in the science modules depends on students being clear about what they need to learn and the teachers knowing what to look for as evidence of students' learning. When the teachers ask themselves how each instructional activity contributes to achieving the learning goal, the aims of the activity, and what students need to pay attention while conducting the activity, becomes clearer, as well as how to know whether students are learning (or practicing) what they have to. Instructional activities should be designed to have a specific purpose, which will allow teachers to easily identify students' various struggles as reflected in the problem-solving strategies they apply. The aim is to prepare teachers to quickly identify the source of student difficulty and provide appropriate assistance.

### ***13.4.3 Design Tasks that Will Demonstrate Whether Students Can Transfer Their Learning***

As mentioned earlier, the mapping process originated as part of a project funded by the National Science Foundation; it helped to build an approach to develop instructionally sensitive assessments (Ruiz-Primo and Li 2008). It was intended to help identify the big, overarching ideas that become the constructs to be tapped with newly developed assessment items.

The information collected from the teachers' unit maps was considered as *sources of instructional sensitivity (SOIS)* that could be manipulated in order to adequately assess students' transfer of learning. For example, knowing the characteristics of the instructional activities in which students were engaged supported the development of items considered close to the curriculum because they were based on the content and the activities described by the curriculum (Ruiz-Primo et al. 2012). These items are considered highly sensitive to the instruction students received. If items yield information proving that students have learned what is expected, this can be a departure point for pushing students to transfer what they have learned (Ruiz-Primo et al. 2012). Without knowing whether initial learning occurred, it is hard to determine whether transfer should be expected (Bransford et al. 1999). In this situation new problems or tasks can be designed as opportunities for students to apply what they learned by manipulating, for example, the context of the items (Ruiz-Primo et al. 2012). The extent to which new problems or tasks differ from the initial learning will determine the 'farness' of the transferability (Barnett and Ceci 2002; Hickey and Pellegrino 2005).

### ***13.4.4 Use the Collected Information: The Assessment-for-Learning Cycle***

A defining characteristic of assessment for learning is that teachers use the collected information to reduce the gap between where students are and where they should be. If teachers do not act on the collected information in a thoughtful way, the gap remains and learning does not improve. Failure to complete the cycle (share learning goals and expectations, collect information, analyze and interpret the information, and *use the information* to develop follow-up strategies to help students to reduce the gap) compromises the potential impact of assessment for learning for improving students' learning.

In an assessment for learning project (Ruiz-Primo and Sands 2010), we coded videos collected during the implementation of 26 entire instructional units in mathematics and/or science with 20 elementary and middle school teachers (Ruiz-Primo et al. 2014). We also conducted observations with 101 teachers in real time (Ruiz-Primo et al. 2016). The results of these experiences confirmed that supportive feedback can be provided only when there is sufficient clarity about the



learning goals, about the evidence needed for determining whether students are learning, and about the best strategies for meeting students' needs. For both informal and formal assessment for learning, we observed how instructional activities can become wasted opportunities when neither the teachers nor their students know what they need to learn. Activities were completed, but opportunities to gather crucial information were lost, as were the opportunities to provide useful feedback or to apply an instructional strategy that would support students' learning.

It is thus continually necessary for teachers to have strong understanding of the 'what', the 'why', and the 'how' of learning, which will enable them to identify the opportunities and the strategies to take advantage of assessment for learning episodes (cycles) that might otherwise be lost.

### 13.5 Conclusions

A key prerequisite for properly implementing assessment for learning is for teachers to have the big picture of what they will engage in with the students over a period of time. This means that they need to understand and keep in mind the main purpose (s) of a unit and each lesson, as well as the activities. They need to see how the various blocks of knowledge fit together at a higher level, which will prepare them to help their students build conceptual frameworks based on principled understandings. A simple but clear indicator of when teachers lack a full, comprehensive picture of what they are teaching is if they have difficulty explaining why certain activities come before and not after others, or why students are carrying out a certain activity. This lack of understanding diminishes the quality of the assessment for learning that can be implemented because the opportunities to gather information and to act upon that information can be easily lost.

This chapter proposed mapping as a strategy that teachers can learn and easily apply to any content they teach. At the same time, it goes far beyond simply learning the mechanics of implementing the units (Giamellaro et al. 2011a). Our research provides empirical evidence that mapping helps teachers to see the forest rather than just the trees. It helps them to construct a more complete picture of the learning path that they will walk along with their students. This includes understanding the sequence of the learning activities and how they contribute to the achievement of learning goals pursued. Many participating teachers in our assessment project considered the mapping process a professional development experience that allowed them to view the unit in a completely different way, even after teaching the same unit for many years. As one teacher wrote to us after a mapping process:

SO MANY THANKS for the opportunity to stretch my brain, grow as a professional, become a better teacher! The curriculum mapping was all I had hoped it would be. I really think that you should consider 'taking this on the road'. It is exactly the type of professional development/work that teachers need to do to ensure student achievement... (January 29, 2012, Science Fifth Grade Teacher)

Successful application of the mapping process relies on *substantial teacher collaboration*. Working together and talking with each other, they develop a richer, more comprehensive understanding of the lessons, how they are connected, and how they contribute to achieving the learning goals. Based on the analysis of the videos collected, we have learned that during the mapping process teachers discovered new ways to understand, articulate, and deliver the units based on their enhanced understanding of the units. We have evidence that asking how each activity contributes to the achievement of the learning goal was a powerful thinking tool that allowed them to evaluate not only the value of the activities but their own content knowledge.

Based on our experiences, and given the nature of the mapping process, unit maps should be built and owned by the teachers. The discussions around the lessons and the exchanges among the teachers in their groups constitute a significant shared learning experience that cannot be replaced by a final product such as the unit map. There is little or no benefit in giving teachers a map built by someone else. Although mapping may seem a costly process, it can be easily presented as a professional development activity in which teachers usually participate every summer before the school year starts. It can take from one to two days, depending on the number of lessons being mapped. Once the strategy is learned, teachers in small working groups in their own schools can readily apply the strategy from unit to unit.

If group mapping is not possible, there are two more options: (1) Individual mapping only. Although its usefulness has yet to be empirically proved, I believe that individual mapping can be valuable at least to help teachers to understand the sequence of the lessons and the activities within those lessons. Missing, of course, is the richness of the teacher conversations. (2) Individual mapping with group discussion. In this approach, teachers first individually map the units on their own. They then meet with their group for comparison and discussion of the individual maps.

Overall, the studies conducted in our research project showed that the mapping process can lead teachers to deeper insights about what they are to teach, why, and the kinds of events and phenomena they must be aware of to help their students achieve the expected learning. Teachers who understand these learning goals and the links between lessons are well positioned to implement more effective assessment for learning practices.

## References

- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learned? A taxonomy of far transfer. *Psychological Bulletin*, *128*(4), 612–637.
- Bransford, J. D., Brown, A. L., Cocking, R. R. (1999). *How people learn: Brain, mind, and school*. National Research Council. Washington, DC: National Academies Press.
- Brophy, J. (1999). Toward a model of the value aspects of motivation in education: Developing appreciation for particular learning domains and activities. *Educational Psychologist*, *34*(2), 75–85.

- Brown, A. L., Kane, M. J., & Echols, C. H. (1986). Young children's mental models determine analogical transfer across problems with common goal structure. *Cognitive Development, 1*, 103–121.
- Dwyer, C. A. (1994). *Development of the knowledge base for the Praxis III. Classroom performance assessment criteria*. Princeton, NJ: Educational Testing Service.
- Eraut, M. (1997). Perspectives on defining 'The learning society'. *Journal of Education Policy, 12*(6), 551–558.
- Giamellaro, M., Lan, M.-C., Ruiz-Primo, M. A., Li, M., & Tasker, T. (2011a). *Mapping science curricula: A method for supporting teachers in the articulation of learning goals*. Paper presented at the AERA annual meeting, New Orleans, LA.
- Giamellaro, M., Lan, M.-C., Ruiz-Primo, M. A., & Li, M. (2011b). *Addressing elementary teacher misconceptions in science and supporting peer learning through curriculum mapping*. Paper presented at the international annual conference of the national association of research in science teaching, Orlando, FL.
- Hickey, D. T., & Pellegrino, J. W. (2005). Theory, level, and function. Three dimensions for understanding transfer and student assessment. In J. P. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 251–293). Greenwich, CN: Information Age Publishing.
- Lan, M.-C., Li, M., Ruiz-Primo, M. A., Wang, T., Giamellaro, M. (2012). *Linking quality of instruction to instructionally sensitive assessments*. Paper presented at the symposium 'Assessing transfer of learning: Instructionally sensitive assessments, curriculum, and instruction', American Educational Research Association Annual Meeting, Vancouver, BC, Canada.
- Lingard, B. (2007). Pedagogies of indifference. *International Journal of Inclusive Education, 11*(3), 245–266.
- Mercer, N., Dawes, L., Wegerif, R., & Sams, S. (2004). Reasoning as a scientist: Ways of helping children to use language to learn science. *British Educational Research Journal, 30*(3), 359–377.
- Ruiz-Primo, M. A. (2011) Informal formative assessment: The role of instructional dialogues in assessing students' learning [Special Issue in Assessment for Learning]. *Studies of Educational Evaluation, 37*(1), 15–24.
- Ruiz-Primo, M. A., & Furtak, E. M. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment, 11*(3–4), 205–235.
- Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching, 44*(1), 57–84.
- Ruiz-Primo, M. A., Furtak, E., Yin, Y., Ayala, C., & Shavelson, R. J. (2010). On the impact of formative assessment on student science learning and motivation. In P. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 139–158). New York: Routledge; Taylor & Francis Group.
- Ruiz-Primo, M. A., Kroog, H., Richey, N., Iverson, L. L., Chrzanowski, A., Shade, C., et al. (2016). *Development, evolution, and adaptation of measures of formative assessment practices: Lessons from the field*. Paper presented at AERA annual meeting, Washington, DC.
- Ruiz-Primo, M. A., Kroog, H., & Sands, D. (2014). *Teacher–student interactions patterns: Looking at teachers with different levels of formative assessment expertise*. Paper presented at AERA Annual Meeting, Chicago, IL.
- Ruiz-Primo, M. A., & Li, M. (2004). On the use of students' science notebooks as an assessment tool. *Studies in Educational Evaluation, 30*, 61–85.
- Ruiz-Primo, M. A., & Li, M. (2008). *Building a methodology for developing and evaluating instructionally sensitive assessments*. Funded proposal submitted to National Science Foundation, Washington, DC: National Science Foundation.
- Ruiz-Primo, M. A., & Li, M. (2009). *Mapping the intended curriculum*. Internal document. Developing and evaluating instructionally sensitive assessments (DEISA) project. University

- of Colorado Denver: School of Education and Human Development, Laboratory of Educational Assessment, Research and Innovation (LEARN).
- Ruiz-Primo, M. A., Li, M., Birkby, E., Edwards, A., Occhipinti, K., Aaron, M., et al. (2013). *Looking at quality of instruction and students' performance: Where do the teachers' questions come from?* Paper presented at the international annual conference of the national association of research in science teaching. San Juan, Puerto Rico.
- Ruiz-Primo, M. A., Li, M., Wills, K., & Giamellaro, M. (2012). *Approach to developing and evaluating assessments at different distance of a curriculum*. Paper presented at the symposium "Approaches to develop, study, and use instructionally sensitive assessment". Paper presented at the national council of measurement in education annual meeting. Vancouver, BC, Canada.
- Ruiz-Primo, M. A., & Sands, D. (2010). *Developing and evaluating measures of formative assessment practices*. Proposal submitted to the Institute of Education Sciences. Washington, DC.
- Stobart, G. (2014). *The expert learner. Challenging the myth of ability*. Berkshire, England: Open University Press.
- Wang, T., Lan, M.-C., Giamellaro, M., Zhao, D. Y., Birkby, E., Ruiz-Primo, M. A., et al. (2012). *Knowledge of learning goals as a navigation tool in curriculum implementation*. Paper presented at the annual meeting of the national association for research in science teaching. Indianapolis, IN.
- William, D. (2011). *Embedded formative assessment*. Bloomington, IN: Solution Tree Press.
- Windschitl, M. (n.d.). *Identifying big ideas in science*. Unpublished manuscript. Seattle, WA: University of Washington at Seattle.

# Chapter 14

## Assessment for Learning: A Framework for Educators' Professional Growth and Evaluation Cycles

Anne Davies, Sandra Herbst and Ann Sherman

**Abstract** Research in the area of classroom assessment for learning (AfL)—in which students are deeply involved in the formative assessment process—is not only extensive, it is also overwhelmingly positive in terms of its impact on student learning and achievement. This chapter focusses on the authors' work with schools and systems where AfL strategies have been deliberately used with adults in support of professional growth and change. The authors provide examples from the perspective of professional growth and evaluation cycles for teachers and school principals. Whether in a school or a large school system, these two Canadian examples illustrate the use of assessment in the service of adult learning, including redefining reliable and valid evidence of adult learning. Experience across multiple schools and school systems has shown that the deliberate alignment of actions from the classroom to the system—particularly in the areas of evaluation and professional growth—positively implicates and impacts everyone's learning.

### 14.1 Introduction

Quality assessment practices, when used thoughtfully, can transform evaluation<sup>1</sup> and professional growth processes of teachers, principals, and others.<sup>2</sup> As we consider the role of assessment in the service of student learning—clearly articulating quality and proficiency, using those descriptions to engage in self-regulation

---

<sup>1</sup>Policy documents in Canada often refer to the term 'teacher supervision', rather than 'teacher evaluation'. However, for the international audience, the latter will be used in this chapter.

<sup>2</sup>Many educational professionals in public school systems (teachers and principals) are unionized across Canada.

---

A. Davies (✉) · S. Herbst  
Connect2learning, Courtenay, Canada  
e-mail: anne@connect2learning.com

A. Sherman  
University of New Brunswick, Fredericton, Canada

and triangulating evidence of learning—we recognize the ways in which these principles can be used effectively to support adult learning. This looks to be a ‘no brainer’, as a Superintendent recently said to us. But if it is a ‘no brainer’, then:

- Why does it seem that the principles of assessment in the service of learning are relegated mostly to the world of the classroom—to be employed by teachers as they teach students?
- Why are parallel expectations and actions across an organization and its roles and responsibilities—alignment—so difficult to systemically and deeply achieve in relation to AfL?
- Why is proof of success often limited to numerical data, rather than being comprised of evidence from multiple sources collected over time?

School systems employ teachers and principals and a wide range of other staff.<sup>3</sup> The employee–employer relationship, while collaborative, is nonetheless hierarchical. Just as classroom teachers are required to evaluate students at times prescribed by policy, leaders, such as principals (who must evaluate teachers) and superintendents (who must evaluate teachers and/or principals), do so at times and in ways prescribed by policy. Policies govern both evaluation and professional growth cycle processes. While both processes support the learning and development of educators, they have distinct purposes. As noted in one school district policy document, evaluation and professional growth are ‘intended to assist teachers in meeting their professional responsibilities and to enhance teaching knowledge, skills and attributes that maximize student learning’ (Edmonton Public Schools 2015, Policy FGCA.AR). The result of both, when done well, is learning.

The evaluation and professional growth cycles are necessarily different. The evaluation cycle, as dictated by policy, is a time of professional appraisal, whereby the employer or supervisor makes a professional judgment regarding the employee’s level of performance. This is just like the classroom teacher who, as dictated by policy, must make a professional judgment of students’ levels of performance on report cards. Professional judgment, in both cases, is informed by one’s knowledge of context, evidence of learning, methods of collecting evidence, and the criteria and standards that describe success. ‘In professional practice, judgement involves a purposeful and systematic thinking process that evolves in terms of accuracy and insight with ongoing reflection and self-correction’ (Ministry of Education of Ontario 2010, p. 152). Both the professional growth and evaluation cycles are a time of learning; however, the former does not require evaluative and summative statements from the supervisor. Yet, both can occur through multiple opportunities and learning pathways that address the uniqueness of each adult learner and enable choice, while affirming a common learning destination.

The balance between the evaluation and professional growth cycles can be delicate—one of supporting learning without judgment and yet, when required,

---

<sup>3</sup>A school system is defined, for example, by its capacity to direct policy development, hire and evaluate staff. In Canada, public school systems vary in size up to 250,000 students (595 schools).

making a judgment, an evaluation. There often seems to be an assumption that the evaluation cycle is 'higher stakes' than the professional growth cycle, yet policies related to professional growth in numerous Canadian jurisdictions such as British Columbia, Alberta, Saskatchewan, Manitoba, and Ontario, clearly articulate a requirement that evidence of professional learning be shown to a supervisor at times and in a manner articulated in policy.

Readers need to note that we are deliberately using the language of schools and systems in order to bridge the understandings of the past with the realities of education today. For example, we are not using 'professional development' and 'professional growth' as synonyms, as the former is most often referred to as a structure or opportunity for learning (e.g., workshop, institute, course, or Professional Learning Community meetings), while professional growth refers to the learning that takes place. 'Evaluation Cycle' is a time dictated in policy where the employer makes a determination regarding the employee's level of performance.

The assumptions that underlie this work in Canada have been changing. In the past, some practices were more typical and now new practices are emerging. Some examples include:

- In the past, teachers evaluated everything. Today, teachers are more likely to evaluate less and spend more time using AfL—formative assessment plus the deep involvement of learners in the assessment process—to support all learners.
- In the past, professional development tended to be more directed and focussed on 'delivery' of information and knowledge whereas now there is more likely to be an emphasis on professional growth—the construction of knowledge through multiple opportunities, varied learning styles, and multiple learning pathways that address the uniqueness of each adult learner and enable choice while affirming a common learning destination.
- In the past, teachers' professional judgment was considered by many to be 'in place' by virtue of qualifying for a teaching credential. These days, 'informed professional judgment' is coming to be viewed as an ongoing learning process that reflects professional knowledge of performance expectations, context, evidence of learning, methods of collecting evidence, and the criterion standards that indicate success.
- In the past, evaluation of educators was often about making a judgment—was the teacher fit to teach? Was the school principal fit to lead? Currently, even if one is on an evaluation cycle, there is a growing tendency for learning to be the expectation and, therefore, AfL principles still apply. While distinct, there is a growing interest in working to ensure evaluation and the professional growth cycles also support teacher learning and development by teachers, principals, superintendents, and others.
- In the past, it was assumed that only the educator involved in the professional growth or evaluation cycle was going to learn. Now, more and more leaders, understand that the learner, the evaluator or supervisor, and the system itself can learn as a result of these processes.

- In the past, numerical data alone seemed to be valued. Now policy in many jurisdictions articulates that evidence of learning must be triangulated, that is, collected from multiple sources—products, conversations, and observation—and collected over time.

Too often in education, system priorities seem to suggest impact merely at the classroom level. ‘To expect only classroom teachers to shift in their work in the absence of systemic realignment is to separate the interdependent parts of the whole’ (Davies et al. 2012b, p. 18). However, recent research suggests, ‘when leaders employ the tenets of AfL as their leadership stance and actions, they exert their leadership in incredibly impactful ways’ (Davies et al. 2014, pp. 588–589). Therefore, the ‘no-brainer’ that is often referred to and consigned to the classroom can, in fact, transform traditional teacher professional growth and evaluation processes. AfL strategies can be deliberately used with adults in support of growth and change in schools and systems if key guidelines related to quality classroom assessment such as triangulated evidence of learning are met. In this chapter, we illustrate that although the purposes for educator professional growth and evaluation cycles may be different (Marzano and Toth 2013), they can, in fact, both be informed by, and use, the principles of AfL.

## 14.2 Research Foundation

Since Black and Wiliam’s (1998) study, AfL has increasingly become the focus of professional learning for teachers. Their research claims that AfL has the greatest impact on student learning and achievement ever documented have served as a catalyst for moving classroom assessment to the centre of the educational agenda. And while their findings have been challenged (Bennett 2011), policy documents continue to acknowledge the importance of formative assessment and involving students in the assessment process (Hawai’i Department of Education 2014; Manitoba Education, Citizenship and Youth 2006<sup>4</sup>; Ministry of Education, Ontario 2010).

Over the past thirty years, classroom assessment has become a recognized field separate from measurement and evaluation (Chappuis et al. 2012; Crooks 1988; Davies 2011; Natriello 1987; Stiggins and Bridgeford 1985). Quality classroom assessment:

---

<sup>4</sup>In December 1993, the ministers responsible for education in Manitoba, Saskatchewan, Alberta, British Columbia, Yukon Territory, and Northwest Territories signed the Western and Northern Canadian Protocol for Collaboration in Basic Education (WNCP), Kindergarten to Grade 12. In February 2000, Nunavut also joined WNCP. In 2006, Manitoba Education, Citizenship, and Youth published the policy statement referred to here about assessment and evaluation on behalf of the WNCP.



- Requires teachers to know and understand the relevant standards/outcomes and agreed-upon statements of quality,
- Uses evidence of learning collected from multiple sources over time,
- Involves AfL to engage the learner and support ongoing learning, and
- Depends upon informed professional judgment rather than external measures (Davies 2011; Davies et al. 2012a, b; Herbst and Davies 2014).

Researchers have shown that when teachers use AfL, students learn more and teaching becomes more effective (Allal 2010; Andrade and Cizek 2010; Andrade 2013). There is also a growing body of research focussed on the power of using AfL in support of adult learning (e.g., Boud et al. 2015; Klenowski and Wyatt-Smith 2013; Sadler 2013). Further, as systems have moved to greater fidelity with the recommendations arising from research related to assessment in the service of learning, researchers have also engaged in examining the system-level implementation process. For example, Gardner (2012) and James et al. (2007) documented implementation across schools and groups of schools in the United Kingdom. Swaffield (2013) and Swaffield and MacBeath (2008) worked with school leaders internationally studying the leadership required for successful implementation of AfL across classrooms and schools. In 2014, we reported on longitudinal research related to using AfL as both the change itself and the way to achieve the change result across a school system (Davies et al. 2014). It documented that positional leaders found more success when they themselves used AfL to support the system-learning initiative. Three of the actions related to AfL that leaders used are relevant to the topics of teacher evaluation and teachers' professional growth:

- Use AfL as a leadership tool (showing samples, co-constructing criteria, coming to common agreement around quality) to do the work they are meant to do,
- Model and coach others using AfL principles, structures, and strategies,
- Use AfL principles, structures, and strategies with every group implicated in the system-learning initiative (students, teachers, administrators, trustees, parents, unions) (Davies et al. 2014).

### 14.3 Two Examples of Educator Professional Growth and Evaluation

In this chapter, we present two examples of educator professional growth and evaluation from different perspectives. One is a system leader—a superintendent—supervising a principal (Manitoba) and the other is a school principal supervising a group of teachers (British Columbia). The actions taken include (Davies 2011):

- Beginning with the end in mind,
- Describing quality,

- Triangulating evidence of learning,
- Engaging the learner in the classroom assessment process,
- Evaluating and reporting the learning.

When district and school leaders apply the practices typical of quality classroom assessment to their work in the area of professional growth and evaluation cycles, they promote alignment. In the following sections, each of the five actions mentioned above is illustrated through two vignettes—one concerning the experience of the superintendent, the other the experience of the principal. These vignettes are based on multiple observations and conversations from our work with school and system leaders (Davies et al. 2014).

### ***14.3.1 Beginning with the End in Mind***

And, as these two examples show, preparation for either the professional growth or evaluation cycle requires precision of purpose and goal; the ‘where we are going’ is necessary to reach success. Taking time to determine just what the teacher and the principal want to learn as a result of professional inquiry, or what the foci of the evaluation process are, is a critical first step. This is no different than in classrooms with students. Whereas both students and teachers are informed through the assessment and evaluation process, the teacher has the professional responsibility to make the final evaluation that is then recorded and reported. Part of the process is being clear about what is to be learned; that is, what knowledge, understanding, application, and articulation need to be shared and demonstrated? When the outcomes are clear, this clarity can later be used to inform the professional judgment of the supervisor against the identified learning outcomes being made.

#### **14.3.1.1 Superintendent**

A principal of five years’ experience began her sixth year with a meeting with the superintendent. It was the year of formal evaluation, as prescribed by policy, and this initial conversation was meant to set the stage for what was to come. Together, they reviewed the school plan results from the past couple of years, the district document outlining the indicators of effective leadership, and the statements of professional growth from the past two years. Consequently, the principal identified a characteristic of leadership that she wished to focus on during the upcoming year (modelling the school’s values and practices) and included outcome statements connected to both teacher and student learning and achievement. More specifically, the latter iterated itself in an increase in the percentage of students who were reading at or above Grade 7 level. From the superintendent’s perspective, these areas of foci aligned with the district priorities; however, he added one additional attribute from the district document—buffering staff from distractions to their

work—and one related to mathematical achievement for students, based on a trend noted in the provincial assessment results.

### **14.3.1.2 Principal**

In one school, the school's learning goal related to the board priority focussed on AfL. All teachers, whether on a professional growth or evaluation cycle, identified their own professional growth plans to learn more about beliefs, attitudes, and practices regarding classroom assessment. The group of teachers in their evaluation year then met with the principal one-on-one. The principal outlined the process, making links to the classroom assessment process explicit. Each teacher was asked to reflect on his/her current classroom assessment practice, reviewing evidence of his/her learning and improvement in relation to the criteria established.

Notice the stakes for showing learning improvements are different, depending on whether the teacher or the principal is on an evaluation cycle or a learning cycle.

## **14.3.2 Describing Quality**

Just as students ask 'What do you want?' or 'How good is good enough?', educators also ask 'What does excellence look like?'. A second similarity in both examples is the need to get to a degree of specificity regarding what quality and proficiency are. Statements of effective teaching, or leadership practice, or district priority statements often define what one should be able to do *without* communicating what it looks like when that is attained. So just as teachers work to look at samples and other data to inform students' expected levels of quality, educators engage in similar processes to more fully understand what is expected of them in terms of 'What does it look like when I learn more about \_\_\_\_\_ in my professional inquiry?' or 'What can a distinguished level of teaching and leading look like in relation to \_\_\_\_\_?'. The responses to these questions serve both the educator and/or the person who is responsible for the evaluation process. It means that there is enough detail and information so that educators can coach themselves and others, regardless of current understanding or performance, towards success.

### **14.3.2.1 Superintendent**

The purpose of the next meeting between the superintendent and the principal was to build a list for each of the two characteristics selected (see below). Certainly, the district's document provided some clarity, but the process of collaboratively describing what each meant garnered greater precision and ownership. The dialogue clearly identified what each person viewed as quality and proficiency. There was no longer room for supposition or assumption.

## **Descriptions of Quality of the Two Leadership Areas of Focus in the Evaluation Cycle**

### **Model the School's Values and Practices**

- Demonstrate sound understanding of current pedagogy and curriculum in reading and mathematics
- Ensure assessment and evaluation practices throughout the school are equitable and appropriate
- Ensure instructional practices use appropriate pedagogy to respond to different needs of learners
- Recognize the potential of new and emerging research in instruction and assessment
- Model professional learning to staff, students, community
- Collaborate during planning cycles
- Analyse a wide range of evidence to determine school progress and growth.

### **Buffer Staff from Distractions**

- Use professional judgment to determine what is brought to staff
- Minimize distractions and disruptions to instructional time
- Engage in collaborative decision making to respond to external requests and initiatives
- Monitor staff participation in out-of-school and non-instructional activities
- Review the ways in which out-of-classroom events and activities contribute to students' learning needs or curricular expectations.

### **14.3.2.2 Principal**

The group of teachers on the professional growth cycle made plans on their own and with each other about their expected learning outcomes and the commensurate student learning outcomes. They also identified types of learning strategies and actions in which they would engage in order to meet their learning goals.

The teaching staff on the evaluation cycle, along with the principal, developed a list of what was important. They were asked to examine their own teaching, learning, and assessment practices and to consider the practices of others; they read professional materials and current research to inform their understanding; they worked as a group to build common understandings. They talked about the important role of assessment and of the role of student evidence in assessing effective teaching—that what is learned is a more important assessment of teaching effectiveness than something merely being taught or 'covered'. They created a comprehensive list with a great diversity of ideas represented. They grouped similar ideas together. Participants worked together to identify criteria by expressing the big idea of each grouping in summary form in an easy-to-understand phrase. The

next step was to list all the possible evidence for each criterion of the learning destination.

Each of the two examples clearly articulates what is to be learned or what is to be present in the teaching or leading. Once the level of quality has been described, the next step is the selection of evidence of learning.

### ***14.3.3 Triangulated Evidence of Learning***

Whether it is learning, progress, growth, or the expected levels of performance that have increased or been met, the identification of evidence is key. When there is a plan to collect evidence of learning from multiple sources over time in relation to what needs to be learned or achieved, the findings are more likely to be reliable and valid than the more limited data sets that have traditionally been used in professional learning or evaluation cycles.

#### **14.3.3.1 Superintendent**

And finally, for each of the leadership areas of focus and the student achievement outcomes, the two jointly created a list of potential evidence that could be collected in order to prove that the characteristics and the student learning outcomes had been met (see below). This conversation was critical—not only to the process but to the commitment of alignment that the school district had made to its staff and partners. This list was no longer one dimensional. Rather, it was to be triangulated at its core. No longer were students judged only on a limited set of evidence and the same was true for the principal. Her performance would be judged on evidence from multiple sources—products, observations, and conversations.

#### **List of Potential Evidence to Be Gathered Related to the Two Areas of Focus in the Evaluation Cycle**

- Discussion regarding analysis of school evidence
- Visual representation of school-based evidence
- Planning notes for collaborative planning cycles
- Staff interviews
- Professional growth plan
- Reflections on professional growth and notes regarding recent pedagogies, research examined
- Pictures and videos of students and teachers at work
- Discussions regarding decision-making processes about what is brought to staff and what is not brought to staff
- Samples of what has been brought to staff and what has not been brought
- Recording notes regarding out-of-classroom activities and events.

### **14.3.3.2 Principal**

Teachers on the professional growth cycle identified triangulated evidence of success that they would use themselves to monitor their progress and could share with their colleagues and principal during biannual ‘check-in’ conversations.

The teachers in the evaluation cycle also considered and listed all possible products, conversations, and observations—anything that could be considered proof of learning. The principal prompted the individual conversations by asking questions such as: ‘What would you see if you spent time observing in the classroom?’, ‘What would you hear?’, ‘What would students say?’, ‘What would parents say?’, ‘What kind of products might be collected?’, ‘What form might they take?’, ‘Who might collect them?’, and ‘When?’. The list of evidence was far more than what could be collected through three formal observations made by the principal. It was obvious that both the teacher being evaluated and the principal doing the evaluation would be engaged in ongoing evidence collection during the year. During the next meeting, the teacher and the principal both brought forward evidence to finalize the baseline evidence collection regarding strengths and goals for improvement.

It is evident that the educators in both examples deliberately planned to collect evidence from multiple sources in relation to the learning focus—an essential aspect of validity.

## ***14.3.4 Learners Active and Engaged During the Process***

In both examples, the learner–educator is the central focus of the process. The educator being evaluated is involved in directing their own next learning steps in relation to the goals set and the descriptions of quality and proficiency established. The educator is deliberately collecting the evidence of learning. In a formal evaluation cycle, the supervisor also gathers evidence.

### **14.3.4.1 Superintendent**

Now that the superintendent and the principal had established a focus for the evaluation cycle, the times they had spent, both together and individually, in that pursuit were focussed and aligned.

The superintendent created a timeline of evidence collection. The evidence, as noted earlier, went beyond observations that were general in nature but would serve the areas of growth that had been identified. These included interactions between staff and students in pursuit of the learning achievement targets. Documents and products were also gathered.

Every other month during the ten-month school year, the superintendent sat with the principal and posed two types of questions. The first centred on what the principal had been doing and learning in relation to the areas of focus. ‘What

actions have you engaged into build your understanding, and what have you done toward the leadership characteristics and the student achievement targets?' 'What new thinking has emerged for you as you have examined the evidence you have been collecting or as you have taken part in both formal and informal learning opportunities?'. The second type of question emphasized the principal's interpretation of the evidence that the superintendent had collected. Examples included: 'What are you noticing about...?', 'What patterns and trends are you seeing as you look through...?'. When reviewing the evidence, the superintendent did not offer his interpretation but rather encouraged the principal to make meaning herself.

The principal carefully considered the criteria for each of the characteristics. As she reflected on her practice, she identified ways that would help her learn more. This included reading professional articles, watching videos of teachers, students, and leaders in action, attending professional learning sessions, and, for her and most importantly, networking with her valued colleagues. She also consciously created a timeline to collect the evidence that had been identified at the outset, including baseline student evidence, professional journal entries, and conversational data from teachers and students. Preparation for regular meetings with the superintendent was minimal, as the focus of these meetings was to reflect on the evidence collected since the previous meeting.

#### **14.3.4.2 Principal**

During the individual biannual meetings between teachers in their professional growth cycle and the principal, the evidence teachers were collecting in relation to their personal learning goals was shared. The principal asked questions to stimulate the conversation, such as, 'What does this evidence tell you about what you are learning?', 'In what ways does your learning support district priorities?', or 'What would you like to learn more about?'. These questions are not evaluative in nature; rather, they consistently turn the learning back to the teacher him- or herself.

After the evidence of learning had been collected from multiple sources, principal and teachers in the evaluation cycle met individually and discussed what the evidence signified. Teachers kept a professional portfolio modelled after the student portfolios focussed on growth over time in relation to the learning goals, showing both beginning evidence and evidence of change. One common self-assessment reflection frame used was, 'I used to... and now I...'. They recorded, either in print or digital media, the changes that had occurred over time in terms of student learning and adult learning. They described the ways their teaching practices had improved. They also identified possible next steps for improvement and set plans for future professional growth. One teacher, having been part of numerous discussions regarding evidence-based professional learning, chose to build the central collection of evidence around three students' work samples. She deliberately selected a student who was excelling, one who was on track to do well, and one student who needed significant learning support. The portfolio initially consisted of baseline collections of student evidence for each student that included observations,

conversations, and products related to literacy and numeracy. At the end of the first term, another collection for each of the same three students was added. The teacher presented an analysis of the growth and development of each of the three students over the term. This process was repeated at the end of each term throughout the school year.

The principal posed questions that allowed each teacher to examine more closely the triangulated evidence that had been collected. Toward the end, the teacher finalized the collection of evidence ensuring it was collected from multiple sources—products, observations, and focussed conversations. The principal and teacher met and reviewed the evidence. The principal asked questions seeking to understand how the evidence showed the teacher’s learning and work towards the goals set at the beginning.

In both cases, the evidence is examined on a regular basis. Consequently, immediate adjustments to next learning steps are made. During the professional growth cycle, the supervisor acts as a coach and a facilitator. During the evaluation cycle, the supervisor is also a coach and a facilitator and, when it is time, is required to make an evaluation and to record and report it to the organization. Notice that the role of coach and facilitator allows for the supervisor to also learn about their role specifically and their role in adult learning; performance management literature refers to this as ‘reverse feedback’.

### ***14.3.5 Evaluating and Reporting the Learning***

In the examples of the professional growth cycle, it is the teachers, themselves, who establishes summary statements of what has been learned and how, while highlighting the evidence of that learning. In the example of the principal and teacher evaluation cycle, the supervisor evaluated the evidence; that is, he/she appraised the evidence with respect to excellence or merit. Each exerted their professional judgment in relation to these questions:

1. What does the adult learner know and what is she/he able to do, and articulate?
2. What areas require further attention or development?
3. In what ways can his/her learning be supported?
4. How is he/she progressing in relation to the set learning goals?

#### **14.3.5.1 Superintendent**

At the end of the evaluation cycle, the final report, though penned by the superintendent, held no surprises for the principal. Because the principal had participated in identifying the areas of focus and collaborated in developing the descriptors of excellence, she knew what was expected. Throughout the year, she was engaged in reviewing evidence collected by another, and she was implicated in gathering



evidence herself. She then reflected on what the evidence pointed to as next possible steps of learning, action, regulation, and intention. As a direct result, the evaluative statements and judgments of her performance and practice as a principal were not based on a narrow band of data. Patterns, trends, and gaps were rooted in the triangulated evidence collected over time.

### **14.3.5.2 Principal**

At the end of the year, the teachers in the professional growth cycle shared end-of-year statements and reflections with each other and with their principal. They summarized the learning that had taken place, identified areas of potential next steps, and reviewed the evidence that they themselves had collected.

For the teachers in the evaluation cycle, the principal reflected on the year's entire collection of evidence, including the notes from the meetings that had occurred over the year. The process of making a professional judgment—the evaluation—was supported through the criteria that had been set earlier.

Decisions were made about how the evidence best demonstrated what had been learned. The principal was able to exert professional judgment with confidence as a result of being engaged in learning, studying district policy and regulations, as well as experiencing a similar process for the principal appraisal process. The series of learning experiences set out by the district to explain and model the teacher evaluation process, including analysing classroom video footage, using the district's revised classroom walk-through framework with colleagues, being mentored while serving as an assistant principal also made this process more likely to be implemented. Because of deliberate alignment, the process this principal followed was the same process district principals used to evaluate the work of district staff and the same process the superintendent used to supervise principals.

What made a quality teacher evaluation report was clear to all because the leadership team at the district level had examined samples of reports and co-constructed criteria. That said, every teacher's evaluation report was a different kind of challenge.

Once the evaluation report was drafted, the principal again sat beside each teacher in the evaluation cycle and reviewed the draft report and the evidence collected. There was an opportunity for the teacher to ask clarifying questions and make suggestions. Then, the principal finalized the report and submitted it to the superintendent.

In this school system, there is an expectation that the adults share their progress, experiences, and results with others in the school community, just as students share evidence of their learning with teachers and parents. So, when the evaluation process was finalized and final meetings had occurred with those in the professional growth cycle, teachers were invited to share evidence of learning as part of the school's collection of evidence. This complete collection was shared with the Board of Trustees as part of the data of school board achievement and success.

Note that in both examples, collections of evidence were reviewed and examined in relation to the initial goal or focus. Together the leader and adult learner, or the adult learner him/herself, reviewed the triangulated collection of evidence collected over time. In the examination of this evidence, they considered ‘best evidence’ in terms of validity and reliability.

## **14.4 From Challenges to Opportunities—Alignment of Purpose and Action**

The origin of the term ‘principal’ was ‘principal teacher’. Principals and superintendents were seen to be teachers of teachers. Recognizing that part of one’s leadership role is that of ‘teacher’ can shift one’s thinking regarding the learning of others (Senge 1990). Researchers have emphasized the importance of school and system leaders understanding AfL and being supportive of its use as a key instructional strategy (Assessment Reform Group 2002; Black et al. 2003; James et al. 2007). Recent teacher evaluation research questions current practices and examines new challenges as a result of calls to use the evaluations in increasingly impactful ways (Bill & Melinda Gates Foundation 2013; Marzano and Toth 2013; Stiggins 2014). And, as research related to positional leaders is beginning to show, the deliberate use of assessment for adult and school learning positively impacts the learning of adults and systems (Davies et al. 2014).

As these two examples illustrate, the actions and strategies of quality assessment most often spoken about in the context of the classroom can be present in the context of educator professional growth and evaluation. This does not occur by chance or as an unintended, yet positive, outcome. Rather, these systems, whether a school or a district, have consciously determined to:

- Describe quality and proficiency,
- Expand proof of success,
- Provide opportunities for learning for all.

### ***14.4.1 Describe Quality and Proficiency***

In the past, there were often no descriptions of quality and proficiency, and the learning focus was not clear to the learners. Principals and other leaders did not always show samples or describe quality. Now clearly defined and agreed upon indicators of quality and proficiency are being developed to bring clarity and transparency not only to student learning but also to the professional growth and evaluation cycles.

### ***14.4.2 Expand Proof of Learning***

In the past, teachers evaluated many specific things and leaders used primarily external scores to determine degrees of success. Today, teachers deliberately evaluate less and spend more time using AfL—formative assessment plus the deep involvement of learners in the assessment process—to support all learners. Now, supervisors also need to learn how to expand proof of learning to support adult learners.

Stiggins (2014) presents an analysis of the kinds of evidence being collected for the purposes of teacher evaluation, including student level data. He summarizes by stating that the evidence typically collected at this point is ‘too thin’. Marzano and Toth (2013) also proposed the evidence being collected should increase in breadth and depth. Darling-Hammond et al. (2012) suggested that teacher evaluation should use professional standards as the source of evidence. They stated:

These standards have become the basis for assessments of teaching that produce ratings that are much more stable than value-added measures. At the same time, these standards incorporate classroom evidence of student learning, and large-scale studies have shown that they can predict teachers’ value-added effectiveness (Wilson et al. 2011), so they have helped ground evaluation in student learning in more stable ways.

We would argue that social science research methods provide a helpful framework for thinking about evidence of learning that is both reliable and valid because it is collected in relation to expectations and standards and arises over time from multiple sources—products, observations, and conversations. This research framework for classroom assessment has a rich history in Canada, dating back to a 1989 curriculum foundation document in British Columbia (Ministry of Education, British Columbia 1989). Triangulating evidence of learning to increase reliability and validity has since gained currency across Canada (Manitoba Education, Citizenship and Youth [WNCY] 2006; Ministry of Education, Ontario 2010).

This classroom assessment perspective acknowledges the complexity of the learning environment and the necessity to collect reliable and valid evidence of learning. All types of learning require evidence of learning that goes beyond common assessments or external measurement data. Rather, quality depends upon the collection and use of a continuous stream of information (both qualitative and quantitative) if feedback is to be specific, if change is to be supported, and if learning is to be successful. This is essential to the inquiry-based nature of successful professional learning at the individual, school, and system levels. Further, research shows that teachers, leaders, and systems learn more (Davies et al. 2014), when leaders:

- Require triangulated evidence of learning from all levels of the system (system, school, appraisal level, and classroom level),
- Transform external pressures (e.g., data from external sources to the school and/or system) into powerful supports for Assessment for Learning goals,

- Value both qualitative and quantitative evidence as proof of student, adult, school, and system learning,
- Model triangulating evidence of learning to inform their own work.

### ***14.4.3 Provide Opportunities for Learning for All***

In the past, evaluation cycles were not codified to include professional growth cycles—they evolved because there was little opportunity for teachers to continue their learning within the organization except under the umbrella of the evaluation cycle. The evaluation cycle tended to occur infrequently and was not sufficient. Now both professional growth cycles and evaluation cycles are more likely to be valued.

Furthermore, in the past, the evaluation cycle was seen to be punitive—something ‘done to others’. Yet, when evaluation takes place in the context of assessment and evaluation that supports learning, adult learners experience the best of classroom assessment and evaluation, i.e., *tight* support and *loose* pressure. Therefore, there is a clear need to help everyone involved in the professional growth and the evaluation cycles understand ways assessment can be used to support the learning of adults, as they both provide opportunities to contribute to everyone’s learning.

In recent years, there has been a clear shift from professional development ‘activities’ to evidence-based professional growth (Darling-Hammond et al. 2012; Guskey 2002; Timperley 2008). This perspective has become more common as illustrated by Mishkind (2014) when she states, ‘High-quality, evidence-based professional development is an ongoing and iterative process grounded in student data. The only real goal of professional learning is to build educator knowledge and skills that will directly impact student learning: their strengths, goals, and instructional needs’ (p. 8).

Teacher evaluation is also an opportunity for leaders to coach, providing specific, descriptive feedback, so teachers find more success. Teacher evaluation is a time when teachers and leaders review professional practice in light of specific goals. Recently, in North America, teacher evaluation has become a ‘hot topic’ in educational circles as initiatives related to merit pay move forward, with many writing about how to do it well (Bill & Melinda Gates Foundation 2013; Darling-Hammond et al. 2012; Marzano and Toth 2013; Stiggins 2014). Often teacher evaluation fails to be viewed in the context of system learning and within the structure that quality assessment provides. This is also true for school and system leaders.

AfL as a transformative tool for schools and school systems is receiving more attention (Davies et al. 2012b; James et al. 2007; Swaffield 2013; Townsend et al. 2010). It has long been acknowledged that schools and systems need to learn (Senge 1990) and systems are composed of people as well as policy, procedures,

regulations, and protocols. It makes sense, therefore, that the learning of adults is also of primary importance for leaders. Leaders are being encouraged to use the practices of assessment for learning to support the learning of adults.

The two examples we have highlighted here show educators' professional growth cycles and evaluation cycles can both be used as an opportunity for system learning as the evaluators, the principal and the superintendent, also learn and, in turn, they can use their experiences to inform the larger school system.

## 14.5 Conclusion

When we first undertook our longitudinal study of professional learning and leadership more than fifteen years ago, we considered educator professional growth and evaluation as an important leadership task and a powerful leadership opportunity. The examples included here illustrate what it can look like in action. Educator professional growth and evaluation cycles benefit from using the principles of quality classroom assessment which are based upon agreed-upon statements of quality, evidence of learning collected from multiple sources over time, and AfL that engages the learner and supports ongoing learning. Further, informed professional judgment in relation to agreed-upon understandings of quality and the valued collaboration between the person being evaluated and the supervisor helps leaders provide tight support in the context of loose pressure. Using these principles of classroom assessment aligns priority, vision, and action across a school system, and as a result, leaders' actions are informed and impactful on student, adult, and system learning.

## References

- Allal, L. (2010). Assessment and the regulation of learning. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (Vol. 3, pp. 348–352). Oxford, UK: Elsevier.
- Andrade, H. (2013). Classroom assessment in the context of learning theory and research. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 17–34). New York: Sage Publications.
- Andrade, H. L., & Cizek, G. J. (2010). *Handbook of formative assessment*. New York: Taylor & Francis.
- Assessment Reform Group. (2002). *Assessment for learning: 10 principles*. Cambridge, UK: University School of Education.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5–25. doi:10.1080/0969594X.2010.513678
- Bill & Melinda Gates Foundation. (2013). *Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three-year study*. Seattle, WA. [http://metproject.org/downloads/MET\\_Ensuring\\_Fair\\_and\\_Reliable\\_Measures\\_Practitioner\\_Brief.pdf](http://metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf). Accessed January 18, 2015.

- Black, P., Harrison, C., Lee, C., Marshall, C. B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Buckingham: Open University Press.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7–75. doi:10.1080/0969595980050102
- Boud, D., Lawson, R., & Thompson, D. G. (2015). The calibration of student judgement through self-assessment: Disruptive effects of assessment patterns. *Higher Education Research & Development*, 34(1), 45–59. doi:10.1080/07294360.2014.934328
- Chappuis, J., Stiggins, R., Chappuis, S., & Arter, J. (2012). *Classroom assessment for student learning: Doing it right—using it well*. Portland, Oregon: Assessment Training Institute Inc.
- Crooks, T. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438–481.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan*, 93(6), 8–15. [http://www.edweek.org/ew/articles/2012/03/01/kappan\\_hammond.html](http://www.edweek.org/ew/articles/2012/03/01/kappan_hammond.html). Accessed January 18, 2015.
- Davies, A. (2011). *Making classroom assessment work* (3rd ed.). Courtenay, BC: Connections Publishing and Bloomington, IN: Solution Tree Press.
- Davies, A., Busick, K., Herbst, S., & Sherman, A. (2014). System leaders using assessment for learning as both the change and the change process: Developing theory from practice. *The Curriculum Journal*, 25(4), 567–592. doi:10.1080/09585176.2014.964276
- Davies, A., Herbst, S., & Parrott Reynolds, B. (2012a). *Leading the way to assessment for learning: A practical guide* (2nd ed.). Courtenay, BC: Connections Publishing and Bloomington, IN: Solution Tree Press.
- Davies, A., Herbst, S., & Parrott Reynolds, B. (2012b). *Transforming schools and systems using assessment: A practical guide* (2nd ed.). Courtenay, BC: Connections Publishing and Bloomington, IN: Solution Tree Press.
- Edmonton Public Schools—Policy Doc #FGCA.AR. <http://www.epsb.ca/ourdistrict/policy/ff/gca-ar/>. Accessed July 3, 2015.
- Gardner, J. (Ed.). (2012). *Assessment and learning* (2nd ed.). London, UK: Sage Publications.
- Guskey, T. (2002). Does it make a difference? *Evaluating professional development*. *Educational Leadership*, 59(6), 45–51.
- Hawai'i Department of Education. (2014). *Educator effectiveness system*. <http://www.hawaiipublicschools.org/TeachingAndLearning/EducatorEffectiveness/EducatorEffectivenessSystem/Pages/home.aspx>. Accessed June 18, 2015.
- Herbst, S., & Davies, A. (2014). *A fresh look at grading and reporting in high schools*. Courtenay, BC: Connections Publishing and Bloomington, IN: Solution Tree Press.
- James, M., McCormick, R., Black, P., Carmichael, P., Drummond, M. J., Fox, A., et al. (2007). *Improving learning how to learn: Classrooms, schools and networks*. London, UK: Routledge.
- Klenowski, V., & Wyatt-Smith, C. M. (2013). *Standards, moderation and teacher judgement: Policy and practice*. London, UK: Sage Publications.
- Manitoba Education, Citizenship and Youth. (2006). *Rethinking classroom assessment with purpose in mind*. Winnipeg, MB: Government of Manitoba.
- Marzano, R., & Toth, M. (2013). *Teacher evaluation that makes a difference: A new model for teacher growth and student achievement*. Alexandria, VA: ASCD.
- Ministry of Education, British Columbia. (1989). *The primary program (Foundation document)*. Victoria, BC: Government of British Columbia.
- Ministry of Education, Ontario. (2010). *Growing success*. Toronto, ON: Government of Ontario.
- Mishkind, A. (2014). *Evidence-based professional learning*. Research Brief #11. California Department of Education. <http://www.calpro-online.org/pubs/evidencebasedprofessionalllearning.pdf>. Accessed January 18, 2015.
- Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist*, 22, 155–175.
- Sadler, D. R. (2013). Opening up feedback: Teaching learners to see. In S. Merry, M. Price, D. Carless, & M. Taras (Eds.), *Reconceptualising feedback in higher education: Developing dialogue with students* (pp. 54–63). London, UK: Routledge (Chapter 5).

- Senge, P. (1990). *The fifth discipline: The art and practice of a learning organization*. New York: Doubleday.
- Stiggins, R. (2014). *Defensible teacher evaluation: Student growth through classroom assessment*. Thousand Oaks, CA: Corwin Press.
- Stiggins, R. J., & Bridgeford, N. J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement*, 22, 271–286.
- Swaffield, S. (2013). Support and challenge for school leaders: Head teachers' perceptions of school improvement partners. *Educational Management Administration & Leadership*, 1–16. doi:10.1177/1741143213494884
- Swaffield, S., & MacBeath, J. (2008). Leadership for learning. In J. MacBeath & N. Dempster (Eds.), *Connecting leadership and learning: Principles for practice* (pp. 32–52). London, UK: Routledge (Chapter 3).
- Timperley, H. (2008). *Teacher professional learning and development*. Geneva, Switzerland: IBE, UNESCO. [www.ibe.unesco.org/fileadmin/user\\_upload/.../EdPractices\\_18.pdf](http://www.ibe.unesco.org/fileadmin/user_upload/.../EdPractices_18.pdf). Accessed March 7, 2015.
- Townsend, D., Adams, P., & White, R. (2010). *Successful assessment for learning projects from AISI cycle 3*. Edmonton, AB: Alberta Education.
- Wilson, M., Hallam, P., Pecheone, R., & Moss, P. (2011). *Investigating the validity of portfolio assessments of beginning teachers: Relationships with student achievement and tests of teacher knowledge*. Berkeley, CA: Berkeley Evaluation, Assessment, and Research Center.

**Part III**  
**Assessment Culture and the**  
**Co-Regulation of Learning**



# Chapter 15

## The Co-Regulation of Student Learning in an Assessment for Learning Culture

Linda Allal

**Abstract** This chapter presents the topic of Part III and the contributions of the chapters included in this Part of the book. The focus of the chapter is on assessment culture, as constructed in classrooms and schools, and on the processes of co-regulation of student learning in relation to assessment for learning (AfL). Co-regulation is defined as the joint influence of student self-regulation and of regulation from other sources (in particular, teachers, peers, assessment procedures, and tools) on the progression of student learning. Summaries of the Part III chapters highlight the new perspectives and the research evidence presented by the authors. A concluding synthesis examines the proposals made by the authors regarding ways of meeting the challenges of AfL implementation. Particular attention is given to professional development and policy measures that are likely to enhance teachers' capacity to implement AfL practices and students' active involvement in these practices.

### 15.1 Introduction

The concept of 'assessment culture' is often mentioned in writings on assessment for learning (AfL) without being clearly defined. In the next section, I propose a definition and cite some examples of research showing how assessment culture is constructed and functions in classrooms and schools. In Sect. 3, I then present a model of 'co-regulation' of student learning in classroom settings and discuss its relationship with the processes and practices involved in assessment for learning.

In the subsequent Sects. 4 and 5, I examine the contributions of the chapters in Part III of this book. Summaries of these chapters highlight the new perspectives that the authors bring to bear on assessment for learning and the range of evidence—from systematic reviews of published research to in-depth studies of classroom practice—that they provide regarding assessment processes and practices.

---

L. Allal (✉)  
University of Geneva, Geneva, Switzerland  
e-mail: linda.allal@unige.ch

The summaries are followed by a synthesis of the authors' proposals concerning the measures needed to move forward with the implementation of high quality assessment for learning.

## 15.2 Assessment Culture

The culture of schooling—and of assessment as a key component of schooling—can have two different meanings. It can refer to a socially transmitted body of beliefs, practices, and tools that are found throughout an educational system. Tyack and Tobin (1994), for example, have described the general 'grammar' of schooling that has prevailed historically in elementary and secondary schools in the United States. Some sociologists (see Anderson-Levitt 2003) go so far as to claim that a worldwide culture of schooling is emerging; they argue that despite important differences in national cultures, schools across the world are becoming more and more similar in their stated goals and modes of organization. With respect to assessment in the classroom, one example is the pervasiveness of the 'I-R-E recitation script' (Mehan 1979): teacher initiates—student(s) respond—teacher evaluates, which has been observed in school systems throughout the world.

A second meaning of culture pertains to the contextualized set of beliefs, practices, and tools that are socially constructed by the actors in each classroom (teachers and students) and in each school (teachers and school leaders, principally). Erickson (1986) has proposed the term 'microculture' to refer to these constructions which are specific to a given teaching–learning community. Although the microcultures of classrooms and schools have common features shaped by system-wide resources (e.g., curricular materials) and constraints (e.g., teaching schedules), they vary considerably with respect to the day-to-day practices of teaching, learning, and assessment, and with respect to the meaning attributed to these practices by teachers and students. It is in this second perspective that assessment culture is considered in this book.

Drawing on insights from research on situated cognition and learning, and in particular the perspective of Cobb et al. (1997), I propose to define assessment culture, as constructed in classrooms and schools, in terms of three interrelated components:

- the *beliefs* of teachers and students regarding the aims of assessment and its relationship to teaching and learning;
- the assessment *practices* that teachers implement and in which students participate;
- the assessment *tools* that support or, in some case, inhibit these practices.

One important concept developed by Cobb et al. (1997) is that meaning (in the present case, the meaning attributed to assessment aims, practices, tools) is never identical in the minds of all actors (students, teachers), but the interactions among

actors allow the emergence of ‘taken-as-shared’ meaning that frames and guides their activity. From this point of view, assessment culture is never completely stabilized. In any given classroom, recurring patterns or configurations can be found, but ongoing interactions always introduce variations and evolutions. To take one concrete example: a teacher may use a standard rubric to provide students with feedback on the narrative texts they produce but may also introduce various forms of differentiation of assessment. For instance, the teacher may initiate a discussion with a particular student in order to understand the difficulties encountered and the next steps to be taken; may suggest that two other students use the rubric to share reciprocal feedback for revision of their texts; may provide still another student who has trouble using standard reference materials with a simplified grammar help-sheet. Although these forms of differentiation are essential to the aim of assessment for learning, students sometimes see them as deviations from the formal equity they consider as a hallmark of assessment. When this occurs, the teacher may need to devise means for constructing with students the taken-as-shared meaning of concepts such as ‘fairness’ versus ‘favoritism’ in assessment (see, for example, Allal 2013).

To summarize, assessment culture needs to be seen as ‘that which weaves together’ (Gallego et al. 2001, p. 960), rather than as a context that surrounds and influences the assessment behavior of teachers and students in a linear manner.

There are relatively few studies that have examined how assessment culture is constructed and functions in classrooms and schools. I will mention two examples that shed light in a significant way on this topic. The first is a study conducted by Mottier Lopez (2005), Mottier Lopez and Allal (2007), over an entire school year, regarding mathematics problem-solving activities in two third-grade classrooms (student aged 8–9) in Switzerland. The focus of the study was on the microculture that emerged in each classroom, principally through the interactions of the teacher and the students during whole-class discussions concerning the problems the students had worked on in small groups. The data collected showed that even when the teachers proposed the same problem-solving tasks to their students and followed very similar pedagogical approaches (alternation between small group work and whole-class discussions), there were differences in the formative assessment dimension of the whole-class discussions. In one classroom, the teacher encouraged students to present their problem-solving procedures in a detailed step-by-step way, guided by teacher scaffolding, but did not actively involve students in the assessment of each other’s proposals; assessment remained the prerogative of the teacher. In the other classroom, when students presented their problem-solving procedures, the teacher often orchestrated peer exchanges during which students expressed opinions about the relevance and effectiveness of other students’ procedures. Although interactive formative assessment, embodied in classroom dialogue, took place in both classrooms and contributed to the regulation of student learning, the assessment microculture differed with respect to the form of student participation.

Research carried out by Birenbaum et al. (2011) addressed the relationship between assessment for learning and assessment culture at both the classroom and school levels. As conceptualized in their study, assessment for learning practices are

embedded in classroom assessment culture, which in turn is nested in the culture of a school-based professional learning community. Culture is seen primarily in ideational terms as the ethos, values, and norms of a learning community. The study included both quantitative analysis of questionnaire data from 122 teachers, working in elementary, middle, and high schools in Israel, and qualitative case studies of six elementary school teachers' AfL practices. The analyses showed that teachers' AfL practices are shaped by several attributes of classroom assessment culture: the epistemological beliefs and conceptions of instruction–learning–assessment, the social climate, and the motivation orientation developed in the classroom. It was found, moreover, that the conception of professional learning in the school was a primary factor that could foster or hinder a classroom assessment culture conducive to assessment for learning.

### 15.3 The Co-Regulation of Student Learning in the Classroom

As indicated in Chap. 1, I propose the following generic definition of the regulation of learning:

Regulation involves four main processes: goal setting, monitoring progress toward the goal, interpretation of feedback derived from monitoring, and adjustment of goal-directed actions and/or of the definition of the goal itself. (Allal 2010, p. 349)

The concept of regulation of learning has been closely associated with classroom assessment in the French-language literature since the late 1970s (see the review by Allal and Mottier Lopez 2005). In particular, a distinction was made (Allal 1979, 1988) between three sorts of regulation stemming from formative assessment: (1) *retroactive regulation*, corresponding to corrective remediation based on the outcomes of formative tests, as proposed in the classical model of mastery learning; (2) *proactive regulation*, involving the use of assessment information to plan new activities adapted to learners' diverse trajectories and interests; and (3) *interactive regulation*, defined as regulation resulting from the learner's interactions with the teacher, with other learners, and with the instructional materials or assessment tools in the learning environment. Interactive regulation is embedded in ongoing learning activities. It highlights learner agency because regulation derives from the (inter) actions undertaken by the learner. The teacher's role is nevertheless important because it is the teacher who sets up the conditions that favor or inhibit interactive regulation.

As conceptualized by Black and Wiliam (2009), formative assessment in support of student learning requires 'the creation of, and capitalization upon, "moments of contingency" in instruction for the purpose of the regulation of learning processes' (p. 6). Retroactive, proactive, and interactive regulations correspond to three different temporal contingencies linking assessment with teaching and learning. Interactive regulation is at the heart of assessment for learning as it takes place in

daily classroom activities, but this form of regulation is never fully effective for every student in every activity. Proactive and retroactive regulations are thus important back-up strategies for adapting instruction to learners' needs and interests.

A distinction has often been made in the psychology of learning literature between regulation that is internal or external to the learner. For example, Brown (1987) introduced a differentiation between 'self-regulation' and 'other-regulation', while in the area of classroom research, Vermunt and Verloop (1999) distinguished 'student regulation' and 'teacher regulation'. The concept of 'co-regulation' offers, I believe, a way of bridging the internal–external distinction. There are, however, several different ways of conceptualizing the co-regulation of student learning.

### 15.3.1 *Conceptions of Co-Regulation*

In a special journal issue on the 'social aspects of self-regulated learning' (edited by Hadwin and Järvelä 2011), the authors discuss ways of enlarging or redefining the field so as to encompass both the self and the social processes involved in regulating learning. The article by Hadwin and Oshige (2011) presents a well-documented analysis that differentiates three categories of regulation of student learning:

- *self-regulation*, which entails active monitoring and regulating of one's own learning;
- *co-regulation*, which is considered as a transitional process in the learner's appropriation of self-regulation strategies through interaction with a more capable other (teacher, more advanced peer);
- *socially shared regulation*, which refers to collective, co-constructed regulation by multiple participants (of equivalent status) who assure the progression of their shared activity.

The above definition of co-regulation is coherent with the approach developed by McCaslin and Hickey (2001) in a Vygotskian perspective on the processes of scaffolding through which the learner internalizes self-regulation strategies. Other authors use the expression co-regulation in a broader sense for any form of socially mediated regulation, whether with a more capable mentor or with peers of equivalent status (Volet et al. 2009).

In my own work (Allal 2007), co-regulation is not conceptualized as one category of regulation among others but rather as an overarching construct that integrates the social and the individual planes of regulation in classroom settings. Co-regulation is defined as the *joint* influence on student learning of the learner's processes of self-regulation and of the sources of regulation in the learning environment: namely, the structure of the teaching/learning situations, the teacher's interventions and interactions with students, the interactions between students, the materials, artifacts and tools used for instruction, and—in particular—for assessment.

This conception of co-regulation is based on two complementary assertions:

1. The first is that *all learning in the classroom is in fact co-regulated*. In other words, I am claiming that self-regulated learning does not exist as an independent entity. Even when students are working alone on a task, actively monitoring their own progress, using strategies to orient and adjust their progression, they are doing so in a context and with tools that are social and cultural constructions. In this perspective, co-regulation is not a transitory process on the way to an autonomous stage of self-regulation. At a given point in time, learners—interacting with teachers and/or peers—do internalize specific strategies of regulation that become part of their repertoire of self-regulation: for example, they may learn to use reverse operations (addition, multiplication) to check the results of subtraction and division. But students in classroom settings immediately encounter new situations (e.g., more complex problem-solving tasks) that require new forms of support (teacher interventions, peer interactions, tools, etc.) in order for them to internalize more advanced forms of self-regulation. Thus, in my view, the learner never ‘slips out’ of the process of co-regulation into a state of autonomous self-regulation.
2. The second assertion is that *processes of self-regulation are nonetheless the core mechanism of learning*. This means that the sources of regulation in the learning environment are simply that—*sources* offering affordances<sup>1</sup> that can enhance or inhibit self-regulation but that do not ensure regulation of learning. The formation of new concepts and the reorganization of conceptual knowledge, as well as the elaboration of learning-oriented motivations, take place through internal processes of self-regulation. In this respect, ‘other regulation’ includes sources of potential regulation (teacher interventions, peer interactions, assessment rubrics, etc.) that have an effect only if they are integrated within the self-regulation system of the learner.

### ***15.3.2 A Model of Co-Regulation and Its Relation with Assessment for Learning***

In the model I have developed (Allal 2007; see Fig. 15.1), the sources of regulation present in the learning environment are nested in a hierarchical manner and the processes of self-regulation are situated at the core of the model. The model includes the following sources of regulation in any teaching/learning situation in the classroom:

---

<sup>1</sup>In resonance with Reed (1996), I propose to define affordances as features (material, technological, cultural, interpersonal) of a learning context that support and at the same time constrain the learner’s activity.

1. the *structure of the teaching/learning situation* (including the learning goals and tasks, the temporal and spatial organization of the situation, the linkages between whole-class, small group and individual phases of work);
2. the *teacher's interventions and interactions with students* (including adjustments the teacher makes to various features of the situation as it unfolds, the teacher's differentiated interventions or forms of scaffolding offered to some students, the teacher's interactions with the whole class, small groups or individuals);
3. the *interactions between students* (including collaborative and cooperative learning activities, as well as peer tutoring).

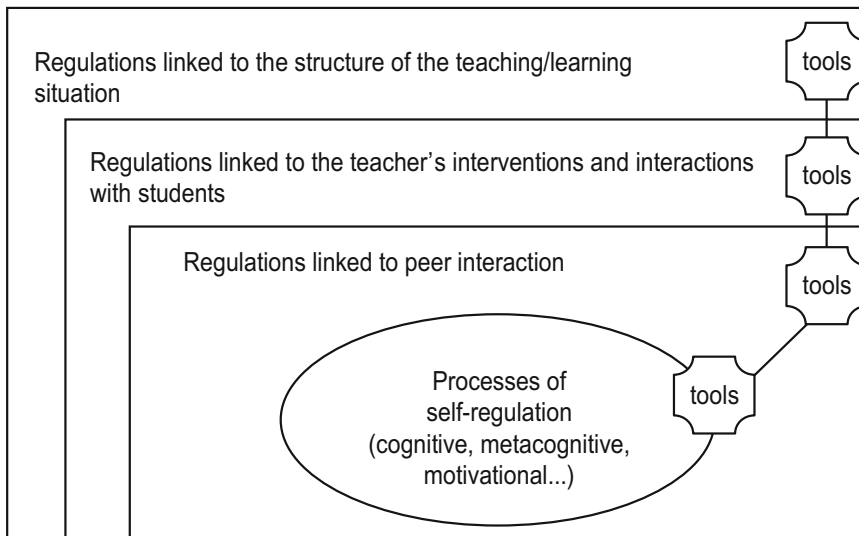
At the core of this nested structure are the processes of self-regulation (cognitive, metacognitive, motivational...) occurring as each learner participates in the teaching/learning situation. Boekaert's (2002) analysis of self-regulation in educational contexts is particularly useful for conceptualizing these processes.

As shown in Fig. 15.1, *regulations linked to tools* are present at each of the preceding levels. Tools include instructional materials, cultural artifacts (e.g., dictionary), technological environments, as well as assessment procedures and instruments. Tools have several important functions in the regulation of learning:

- tools *assure linkages* between the different levels of regulation; for example, a rubric selected by the teacher from curricular materials may then be the focus of a whole-class discussion where more specific success criteria are defined; the resulting extended rubric can then be used by students both for peer assessment and for self-assessment;
- tools *amplify* the effects of interactive co-regulation by making the goals more explicit, the monitoring more systematic, the decision making more evidenced based, than would be the case if the interactions took place without tools;
- tools often allow *recording of traces* of assessment that can be used for deferred regulation; for example, problems which students noted in a group assessment report, and which could not be solved by direct teacher intervention, can become the basis for a new classroom activity.

There are two major implications of the model I propose for assessment for learning. The first is that assessment for learning needs to promote regulations operating—simultaneously or successively—*at several different levels*. For example, self-assessment—a key AfL practice aimed at fostering self-regulation (Andrade 2010)—should not be seen as an encapsulated event: self-assessment operates in a situation with specific affordances; teachers set guidelines and intervene in the students' implementation of self-assessment; even when peers do not directly intervene, they often have an indirect influence (e.g., a student glances at how his or her neighbor is filling out a self-assessment questionnaire...); tools play a critical role (either supporting or inhibiting the student's engagement in self-assessment).

A second implication concerns the *negotiation of meaning* in an assessment for learning culture. The meaning attributed to assessment in general, or to a specific



**Fig. 15.1** Co-regulation of student learning in classroom settings. *Source* Allal (2007), reproduced with permission from De Boeck Supérieur S.A., translated by author

assessment practice, results from the transactions occurring at all the levels in the proposed model. How students understand an assessment, and what they feel about it, depends on the structure of the situation (e.g., is time planned for carrying out adjustments or revisions?), the teacher's interventions and interactions with students (e.g., does the teacher encourage reflection about learning processes?), the interactions with other students (e.g., do they provide constructive feedback?), and the tools used (e.g., do they support learning in an effective, productive manner?). Interactive forms of assessment are occasions for the elaboration of taken-as-shared meaning by the members of a classroom community. As described by Ruiz-Primo (2011), this occurs in 'instructional dialogues' in the context of 'informal formative assessment' integrated in everyday classroom activities. It also occurs in one-on-one teacher-student dialogues that allow scaffolding of learning in the student's zone of proximal development (Allal and Pelgrims Ducrey 2000) and in which the mode of 'teacher questioning' plays a critical role in shaping the dialogue and in promoting student engagement (Heritage and Heritage 2013).

One additional comment on the model presented in Fig. 15.1 which focuses on the co-regulation of learning in classroom settings. Although space does not permit further development here, it is obvious that co-regulation in the classroom is influenced by phenomena outside the classroom, such as the interactions among professionals (teachers, school leaders, etc.) within a school or a network of schools, the resources provided and the policies adopted by the school system, and the long-standing traditions of education in the surrounding community, region, or country.



## 15.4 New Perspectives on Assessment for Learning in Classrooms and Schools

The Part III chapters provide new perspectives on the ways teachers and students engage in the processes and practices of assessment. Part III opens with a chapter by Birenbaum (Chap. 16) that extends her previous work on assessment culture. The principal goal of the chapter is to contrast an ‘assessment culture’ aimed at supporting student learning with a ‘testing culture’ focused on increasing student performance on summative tests constructed by teachers or by the school system. The author bases this contrast on extensive qualitative data (interviews, observations, analyses of artifacts) from four schools, two of which illustrate each type of culture. The analysis describes the mindsets characteristic of each culture and the relationships between student learning in the classroom, teacher professional learning, and school leadership, seen as three interlocking systems. One important implication of Birenbaum’s analysis is that the implementation of assessment for learning must be embodied, in a coordinated manner, in all three systems. She provides a number of concrete suggestions about how to support schools in their transition from a culture focused on testing to a culture oriented toward assessment for learning. In particular she states that professional development in assessment for learning ‘should be contextualized and tailored to a staff’s level of functioning as a school-based professional learning community’ (see Chap. 16).

The next two chapters in Part III (Chaps. 17 and 18) adopt similar approaches: starting with a definition of the processes of self-regulated learning (SRL), they examine the ways in which practices of assessment can foster these processes and thereby contribute to the progression of student learning. Each chapter emphasizes, however, specific aspects of these processes. In Chap. 17, Andrade and Brookhart analyze the relationship between SRL processes (such as forethought, monitoring, reaction, and reflection) and the cyclical phases of classroom assessment, namely: (1) goal setting, (2) progress monitoring, and (3) revision and adjustment. Their analysis is based on a review of published research evidence from multiple sources, including their own extensive studies of assessment processes and practices. The authors’ focus is on *classroom assessment*, which includes formative assessment activities as well as feedback from some forms of summative assessment that may provide support for student learning. They emphasize student agency in assessment but are attentive to the important role of teachers in creating the conditions that allow student involvement in assessment. Formative assessment, especially when it includes self-assessment and peer assessment, is seen ‘as a form of SRL instruction’ (see Chap. 17). The authors provide a number of suggestions of professional development activities designed to promote practices coherent with the goals of assessment for learning.

The authors of Chap. 18 (Panadero et al.) analyze assessment for learning in relation to three cyclical phases of self-regulated learning (forethought, performance, self-reflection). They review published research evidence, from their own work and from a range of other sources, regarding two key AfL practices—student

self-assessment and peer assessment—and the contributions of these practices to self-regulated and co-regulated learning. They emphasize in particular the differential effects on SRL processes of two formative assessment tools: rubrics and scripts. Their chapter includes a series of guidelines for implementing self-assessment and peer assessment in classroom settings, as well as a discussion of the risks and pitfalls associated with these practices, especially in the context of summative assessment. Panadero et al.'s analysis of 'teachers' mediating and modeling role in peer and self-assessment' (see Chap. 18) highlights the joint influence of external and internal processes of regulation on student learning, and in this respect supports the idea that AfL promotes co-regulation.

In Chap. 19, Heritage focuses on student–teacher interactions as a primary source of co-regulation of learning in the classroom. From video recordings collected in a third-grade and a fifth-grade classroom, the author presents and analyzes extensive excerpts showing the contributions of the teacher and of the student to the unfolding dialogue and the aspects of informal formative assessment present in their exchange. Heritage's in-depth analysis highlights key conditions for effective co-regulation of student learning: in the two cases presented, 'the interactions reflected a partnership between teacher and student, where student agency in the learning and assessment process was both acknowledged and supported by the teacher in the context of the classroom's assessment culture' (see Chap. 19). The author believes that teachers need to acquire, as of their preservice preparation, an understanding of the 'social ecology of the classroom' so that they can develop the practices of questioning and interacting with students that promote productive assessment for learning.

The final chapter in Part III (Chap. 20 by Bourgeois) presents an intervention research project designed to help teachers develop AfL practices that promote active student involvement in the co-regulation of classroom learning. The interventions include: teachers' formulation of professional learning goals in reference to the school system's assessment policy; the exchange between teachers of student work samples and the discussion of their respective assessments; the use of a booklet that guides teachers in a series of activities intended to develop student involvement in assessment (including co-construction of criteria, self-assessment, peer assessment). From six case studies of grade 7–8 teachers, Bourgeois presents the contrasting 'assessment journeys' of two teachers and their classes over an entire school year. One teacher finds it difficult to develop new formative assessment practices because she feels overwhelmed by the demands of summative assessment imposed by the school system. The other teacher progresses in her willingness to develop co-regulated formative assessment practices, thanks in large part to the positive experience she and her students have using the proposed booklets. As noted by Bourgeois, in reference to Fullan (2001): this seems to be an instance where 'behavior changes before beliefs' (see Chap. 20). The implications for implementing assessment for learning are discussed.

Taken together, the Part III chapters provide a perspective on assessment for learning that is coherent with the model of co-regulation presented in Fig. 15.1. The authors of each chapter consider students' self-regulation to be the mechanism by which their learning progresses, while at the same time pointing to factors in the

learning environment—the tasks proposed by the teacher, the teacher’s interactions with the students, the role given to peer interaction, the assessment tools utilized—that can enhance or inhibit students’ engagement in the processes of self-regulation that ensure learning.

## 15.5 Moving Forward

This section presents a synthesis of the Part III authors’ suggestions and proposals regarding ways of meeting the challenges of assessment for learning. It is obvious that lessons learned in one context cannot be directly applied in another and that adaptations are always needed. It is nevertheless important to identify the convergent observations and ideas that can provide a basis for moving forward in terms of reflection and action aimed at implementing assessment for learning.

The Part III authors’ observations and analyses of assessment, as practiced in classrooms and schools, or as studied in research projects, lead them to make suggestions and recommendations related to the topics dealt with in the first two Parts of this book: namely, professional development and educational policy in the area of assessment. The main concern is how to design professional development activities and policies that will enhance teacher agency in the development of effective AfL practices and student agency in the participation in these practices.

Most of the authors’ suggestions or recommendations concern the *design of professional development* (PD) activities that will help teachers and other professionals implement effective practices of assessment for learning. The idea of PD activities geared to the zone of proximal development of the actors was already expressed in the introduction to Part II (Chap. 8—Laveault). Part III authors point out that there are sizeable variations among teachers with respect to their existing assessment beliefs and practices (Chap. 20—Bourgeois), as well as important variations among schools with respect to their prevailing assessment cultures (Chap. 16—Birenbaum). This means that professional development activities need to be conducted within a framework that can accommodate several *different zones of proximal development*. Such a framework would allow differentiation of PD approaches in order to take into account teachers’ existing practices and respect teachers’ agency as professionals. It is also proposed that PD be tailored to the needs of each school and may sometimes need to start with professional development for principals or other school leaders before moving on to PD activities for classroom teachers (Chap. 16—Birenbaum).

Regarding the *content* of professional development activities, the Part III authors formulate a number of suggestions. In order for teachers and school leaders to develop deep understanding of the ‘spirit’ of AfL, it is important that PD include reflection about the purpose, rationale, and underlying values of AfL (Chap. 17—Birenbaum). It also needs to address the conceptions of teaching and learning that underpin AfL and the interactive processes that are involved in the co-regulation of student learning (Chap. 19—Heritage).

Key topics to be dealt with in professional development activities include:

- the conceptualization of success criteria that allow students to understand learning goals and to actively monitor their progress toward these goals (Chap. 17—Andrade and Brookhart);
- the formulation of feedback that will be effective for supporting student learning (process-oriented feedback, feedback on processes of self-regulation), as well as procedures for helping students learn how to interpret and use feedback (Chap. 17—Andrade and Brookhart; Chap. 18—Panadero et al.);
- the development of assessment criteria, rubrics, and scripts, through interaction with students in order to increase their capacity to use these references effectively for self- and peer assessment (Chap. 17—Andrade and Brookhart; Chap. 18—Panadero et al.; Chap. 20—Bourgeois);
- the implementation of guidelines for student self-assessment and peer assessment, including analysis of the risks and pitfalls associated with these procedures (Chap. 18—Panadero et al.);
- the development of skills in questioning students and in interacting with students, as well as skills in classroom management that create space for individualized, one-on-one dialogues (Chap. 19—Heritage).

Regarding the *forms* of professional development, the Part III authors advocate different ways of assuring ‘close-to-practice learning’ (Chap. 19—Heritage). Several authors state that PD should be school-based in order to promote a sustainable professional learning community (Chap. 16—Birenbaum, Chap. 18—Panadero et al.). The proposals emphasize collaboration among participating teachers and school leaders in ‘hands-on’ activities such as:

- collaborative design of assessment tasks and tools, including rubrics and scripts, (Chap. 16—Birenbaum, Chap. 17—Andrade and Brookhart, Chap. 18—Panadero et al., Chap. 20—Bourgeois)
- collaborative analysis of student work samples and application of assessment criteria (Chap. 16—Birenbaum, Chap. 17—Andrade and Brookhart);
- collaborative analysis of examples of feedback and production of examples of more effective feedback (Chap. 16—Birenbaum, Chap. 17—Andrade and Brookhart);
- collaborative analysis of samples of teacher–student interactions and of samples of teachers’ practices of questioning (Chap. 19—Heritage).

In addition, several authors observe that if, in the context of a PD activity, teachers have a positive experience with a new assessment practice (such as trying out a peer assessment procedure in their classroom), this experience may have a high potential for transforming their beliefs about assessment for learning (Chap. 18—Panadero et al., Chap. 20—Bourgeois). Finally, the organization of professional development needs to allow teachers and other professionals to experience co-regulation of their *own* learning (through interactions with colleagues, with school leaders, with PD providers or researchers) in order to better

understand how co-regulation may function in the classroom (Chap. 20—Bourgeois). This implies that PD providers engage in modeling the assessment practices they expect teachers to develop in the classroom (Chap. 17—Andrade and Brookhart).

Some of the authors of the Part III chapters also address implications of their research for the formulation of *assessment policy*. Since the practices mandated by a school system's assessment policy are taken up and implemented by teachers in different ways, this needs to be explicitly taken into account in the formulation of assessment policy (Chap. 20—Bourgeois). More generally, policy makers need to realize that their decisions regarding the frequency and intensity of high-stakes testing, as well as the requirements imposed on teachers with respect to summative grading and reporting, determine to a large extent the potential space for developing the formative function of assessment for learning (Chap. 17—Andrade and Brookhart, Chap. 20—Bourgeois). One further suggestion (Chap. 19—Heritage) is to include AfL practices based on teacher–student dialogue within the policy frameworks of teacher evaluation, as a way of recognizing and supporting the key role of these practices in supporting student learning.

## 15.6 A Concluding Reflection

The research conducted by the authors of the Part III chapters, as well as their personal involvement in professional development activities, has led them to a series of observations and recommendations that are globally coherent with those presented in Parts I and II. Effective implementation of assessment for learning clearly requires the *concerted coordination of policy, professional development, and practice in classrooms and schools*. Nevertheless, at any point in time in a given educational jurisdiction, this coordination may be more or less satisfactory. It is therefore important, I believe, that the actors in each arena (policy makers, PD providers, teachers, and school leaders) think about the initiatives that can be concretely undertaken—here and now—in order to move forward with the implementation of AfL. Here are some examples. Policy makers can develop a framework that provides gradual, positive support for assessment for learning even if existing classroom practices are quite far from the goals of AfL. Professional development providers can include principles and practices of AfL in the activities proposed to teachers and school leaders even if, as is the case in many jurisdictions, there is no official policy mandating assessment for learning. Classroom teachers can develop some forms of assessment for learning even when the official assessment policy provides little direct support: for example, if the frequency of summative reporting leaves little space for AfL, teachers may still be able to integrate informal, interactive regulation of student learning within ongoing instructional activities. All of these limited initiatives will not allow optimal implementation of assessment for learning. But they can be first steps of an engagement in this direction. The responsibility of each category of actors should not be contingent on

the others acting first. This would be a recipe for inaction. Each member of the education community needs to act in his or her own arena, while seeking opportunities for coordination with other professionals.

The challenges of implementation of assessment for learning are multiple, and the ways of meeting the challenges are equally diverse, as attested by the chapters in this book. The goal, however, is imperative: namely to ensure that learners develop the knowledge, skills, capacities of reflection, and motivations needed for productive participation in the world of today and of tomorrow. I believe we must all move forward for their sake.

## References

- Allal, L. (1979). Stratégies d'évaluation formative: Conceptions psycho-pédagogiques et modalités d'application. In L. Allal, J. Cardinet, & P. Perrenoud (Eds.), *L'évaluation formative dans un enseignement différencié* (pp. 130–156). Bern: Lang.
- Allal, L. (1988). Vers un élargissement de la pédagogie de maîtrise: Processus de régulation interactive, rétroactive et proactive. In M. Huberman (Ed.), *Assurer la réussite des apprentissages scolaires: Les propositions de la pédagogie de la maîtrise* (pp. 86–126). Paris: Delachaux et Niestlé.
- Allal, L. (2007). Régulation des apprentissages: Orientations conceptuelles pour la recherche et la pratique en éducation. In L. Allal & L. Mottier Lopez (Eds.), *Régulation des apprentissages en situation scolaire et en formation* (pp. 7–23). Bruxelles: De Boeck.
- Allal, L. (2010). Assessment and the regulation of learning. In P. Peterson, E. Baker, & B. McGraw (Eds.), *International encyclopedia of education* (Vol. 3, pp. 348–352). Oxford: Elsevier.
- Allal, L. (2013). Teachers' professional judgement in assessment: A cognitive act and a socially situated practice. *Assessment in Education: Principles, Policy & Practice*, 20, 20–34.
- Allal, L., & Mottier Lopez, L. (2005). Formative assessment of learning: A review of publications in French. In OECD (Ed.), *Formative assessment: Improving learning in secondary classrooms* (pp. 241–264). Paris: OECD.
- Allal, L., & Pelgrims Ducrey, G. (2000). Assessment *of*—or *in*—the zone of proximal development. *Learning and Instruction*, 10, 137–152.
- Anderson-Levitt, K. (Ed.). (2003). *Local meanings, global schooling: Anthropology and world culture theory*. New York: Palgrave Macmillan.
- Andrade, H. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. Andrade & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 90–105). New York: Routledge.
- Birenbaum, M., Kimron, H., & Shilton, H. (2011). Nested contexts that shape assessment for learning: School-based professional learning community and classroom culture. *Studies in Educational Evaluation*, 37, 35–48.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21, 5–31.
- Boekaerts, M. (2002). Bringing about change in the classroom: Strengths and weaknesses of the self-regulated learning approach. *Learning and Instruction*, 12, 589–604.
- Brown, A. L. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 65–116). Hillsdale, NJ: Lawrence Erlbaum.

- Cobb, P., Gravemeijer, K., Yackel, E., McClain, K., & Whitenack, J. (1997). Mathematizing and symbolizing: The emergence of chains of signification in one first-grade classroom. In D. Kirshner & J. A. Whitson (Eds.), *Situated cognition: Social, semiotic, and psychological perspectives* (pp. 151–233). Mahwah, NJ: Lawrence Erlbaum.
- Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 119–161). New York: Macmillan.
- Fullan, M. (2001). *Leading in a culture of change*. San Francisco: Jossey-Bass.
- Gallego, M., Cole, M., & The Laboratory of Comparative Human Cognition. (2001). Classroom cultures and cultures in the classroom. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 951–997). Washington, DC: American Educational Research Association.
- Hadwin, A. & Järvelä, S. (Eds.). (2011). Special issue on social aspects of self-regulated learning: Where social and self meet in the strategic regulation of learning. *Teachers College Record*, 113.
- Hadwin, A., & Oshige, M. (2011). Self-regulation, coregulation, and socially shared regulation: Exploring perspectives of social in self-regulated learning theory. *Teachers College Record*, 113, 240–264.
- Heritage, M., & Heritage, J. (2013). Teacher questioning: The epicenter of instruction and assessment. *Applied Measurement in Education*, 26, 176–190.
- McCaslin, M., & Hickey, D. T. (2001). Self-regulated learning and academic achievement: A Vygotskian view. In B. Zimmerman & D. Schunk (Eds.), *Self-regulated learning and academic achievement: Theory, research, and practice* (pp. 227–252). Mahwah, NJ: Lawrence Erlbaum.
- Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Cambridge, MA: Harvard University Press.
- Mottier Lopez, L. (2005). *Co-constitution de la microculture de classe dans une perspective située: Etude d'activités de résolution de problèmes mathématiques en troisième année primaire*. Doctoral dissertation in sciences of education. University of Geneva.
- Mottier Lopez, L., & Allal, L. (2007). Sociomathematical norms and the regulation of problem solving in the classroom. *International Journal of Educational Research*, 46, 252–265.
- Reed, E. S. (1996). *Encountering the world: Toward an ecological psychology*. New York: Oxford University Press.
- Ruiz-Primo, M. A. (2011). Informal formative assessment: The role of instructional dialogues in assessing students' learning. *Studies in Educational Evaluation*, 37, 15–24.
- Tyack, D., & Tobin, W. (1994). The 'grammar' of schooling: Why has it been so hard to change? *American Educational Research Journal*, 31, 453–479.
- Vermunt, J. D., & Verloop, N. (1999). Congruence and friction between learning and teaching. *Learning and Instruction*, 9, 257–280.
- Volet, S., Vaurus, M., & Salonen, P. (2009). Self- and social regulation in learning contexts: An integrative perspective. *Educational Psychologist*, 44, 215–226.

# Chapter 16

## Assessment Culture Versus Testing Culture: The Impact on Assessment for Learning

Menucha Birenbaum

**Abstract** In this chapter, I contrast two school cultures—assessment culture (AC) and testing culture (TC)—and examine their potential to support or inhibit assessment for learning. First, I compare the two cultures with regard to the mindset that each culture reflects and then with regard to classroom learning, teacher professional learning, and leadership, addressing their interrelationships and their effects on formative school-based assessment. I further compare the impact of external accountability tests on AC and TC schools. In the discussion, I contrast the schools with regard to an underlying dimension of internal coherence, addressing the conditions that foster and those that hinder such coherence. I then explain why AC, as opposed to TC, is conducive to the successful implementation of assessment for learning, and identify the challenges in transforming a TC into an AC. I conclude with recommendations to guide such a transformation.

### 16.1 Introduction

The capacity of assessment for learning (AfL) to promote learning has long been acknowledged (Black and Wiliam 1998). However, judging from AfL implementation in various countries, it seems to have often failed to fulfill expectations, seemingly due to being enacted in ways that conform only to its ‘letter’ rather than embracing its ‘spirit’ (Marshall and Drummond 2006). School culture<sup>1</sup> seems to be an appropriate context in which to search for that ‘missing spirit.’ Through an assessment lens, the school culture continuum ranges from a testing culture (TC) at

---

<sup>1</sup>School culture refers to ‘the underground stream of norms, values, beliefs, traditions, and rituals that has built up over time as people work together, solve problems, and confront challenges’ (Peterson and Deal 1998, p. 28).

---

M. Birenbaum (✉)  
Tel Aviv University, Tel Aviv, Israel  
e-mail: biren@post.tau.ac.il



one end to an assessment culture (AC) at the other (Birenbaum 1996). The purpose of the current chapter is to compare the two cultures and their impact on AFL.

In an earlier paper (Birenbaum 2014), AC was conceptualized as a complex system (Mitleton-Kelly 2003) in which two complex learning systems—classroom learning (CL) and teacher professional learning (TPL)—are nested. It was argued that, through recursive interactions, the three systems influence each other in a co-evolutionary process, which results in the emergence of new knowledge and paradigms. The paper depicted the nature of AC and the mindset<sup>2</sup> that it reflects, characterizing it according to seven categories of assessment-related beliefs and conceptions, which are:

- It's all about learning;
- Assessment drives learning;
- Assessment means dialogue (interaction) with the learner;
- Assessment empowers the learner;
- Diversity is desirable;
- I/we can do it;
- Assessment requires modesty.

For the current chapter, three targets are set. The first is to characterize the TC mindset and to contrast it with the AC mindset. It has been argued that identifying an organization's mindset is relatively easy (Murphy and Dweck 2010); however, understanding its origins is a challenging task. The second target is thus to address this challenge with respect to the two mindsets. To this end, AC and TC schools are compared with regard to classroom learning, teacher professional learning, and school leadership, as well as the impact of external accountability on these cultures. The third target is to discern the relationships among the systems within each culture and identify an underlying dimension that accounts for the differences between the two cultures. Once these targets are achieved, it should become obvious which school culture is conducive for the successful implementation of AFL.

## 16.2 Method

### 16.2.1 *Sample*

This chapter integrates the results from eight case studies of Israeli schools: two pairs of elementary schools (students aged 6–12) and two pairs of middle schools (students aged 12–15). The schools in each pair have a similar demographic profile but differ in the mindset that is reflected in their cultures. The case studies that were

---

<sup>2</sup>A mindset is an implicit theory or a set of beliefs held by people that influences their feelings, choices, behaviors, and outcomes (Dweck 2006).

selected for this chapter are a subset of a larger sample of 42 case studies that were conducted by our research group in the past six years in elementary and middle schools. All of the studies were focused on classroom assessment and school-based professional development (PD) and employed the same data collection techniques. In order to analyze the dominant features of the two contrasting school cultures, half of the schools included in the current chapter were identified, based on the evidence that we collected, as the most representative in the entire sample of an AC mindset and the other half as the least representative of that mindset. Each group comprised two elementary and two middle schools; they will be henceforth referred to as A1–A4 for the first group and T1–T4 for the second group. The elementary schools are A1, T1 and A2, T2; the first two serve a middle-class suburban population, and the other two serve a lower-class inner city population. The other four schools—A3, T3 and A4, T4—are middle schools within six-year secondary schools; the first two belong to the rural education sector serving middle-to-lower-class populations, and the other two are inner city schools serving a lower-class population.

## ***16.2.2 Data Sources and Analysis***

The primary data sources that were used in this chapter include transcripts from semi-structured interviews with 20 teachers, eight principals, and eight focus groups of students. The other sources that were used to corroborate the conclusions from the primary sources consist of transcripts from pedagogical meetings and classroom discourse, assessment-related artifacts (such as tests, quizzes, performance tasks, rubrics, exemplars of written self-assessment, and feedback comments that were given by the teacher and by peers, where available), and school documents (such as school vision, work plans, activity reports, assessment reports, and protocols of pedagogical meetings). Content analysis (Krippendorff 2004) was employed to analyze the evidence.

## **16.3 Results**

### ***16.3.1 TC Versus AC Mindsets***

The content analysis that was conducted on the evidence from interviews with teachers, students, and principals of schools T1–T4 regarding assessment-related beliefs and conceptions yielded seven major categories. Each category, along with an illustrating quote from the interviews, is presented below. Each category is then contrasted with a matched category of the AC mindset that is also illustrated by quotes from the current sample (schools A1–A4). The categories for both mindsets are matched according to the following indicators: purpose of assessment; function

of assessment; method of assessment; power relations in assessment; attitudes toward diversity; expectations about learning; fidelity of assessment.

**The purpose of assessment.** The holders of a TC mindset believe that ‘assessment is all about making the grade.’ An excerpt from an interview with ninth graders reflects this belief:

Student (T): Our school is considered quite an advanced school; most of the students graduate with a good certificate...

Student (U): A high-level school...

Student (T): People hear the school’s name and say, ‘Ah, this is the school of the smart.’ [However] [I]t’s a grade factory; does this mean a good school?

...

Student (M): What they [the teachers] expect of us is that we get high test scores.

Student (T): These are also your expectations for yourself, to succeed in tests.

Student (M): True!

Student (T): And get a better grade, because the rest does not matter.

Student (M): Because, that’s what counts in the end

Student (G): Yes, it is the certificate.

(Ninth graders, School T3)

By contrast, the holders of an AC mindset believe that ‘assessment is all about learning’:

For me assessment is to see where the student is [in his learning] and how I bring him up from there... how I improve... advance him further.

(Teacher, School A2)

**The function of assessment.** The holders of a TC mindset believe that assessment is mainly for accountability:

As a teacher, I feel that they [the tests] examine me more than they examine the students, because at the end, the one who gets the rebuke is the teacher who did not cover the curriculum or did not meet or strengthen areas in which the class failed... (Teacher, School T2)

By contrast, the holders of an AC mindset believe that assessment drives the teaching and learning of both students and teachers:

Assessment not only regards the kids it also constantly has an impact on me because teaching is not a static profession. Each time... even if I teach the same topic it takes on a different shade from class to class, from student to student.... Somehow I do the work... of bettering my profession through assessment.

(Teacher, School A3)

**The methods of assessment.** The holders of a TC mindset believe that standardization is a desirable feature of assessment: ‘I see assessment as objective judgment of student performance according to standards...’ (Teacher, School T1).

By comparison, the holders of an AC mindset believe that assessment means dialogue (interaction) with the learner:

For me, a task or a test is a platform for dialogue with the student. I check the test and I imagine talking with them while I am reading what they wrote. I write them notes in the margins, spontaneous comments that cross my mind as I read. I strengthen, give feedback,

say what's missing and always at the end of each question I actually write a summary that tells the kid what are the good things that he should preserve, what is missing, and how to fix things...

(Teacher, School A3)

**Power relations in assessment.** The holders of a TC mindset believe that the assessor knows best, hence they undervalue students' participation in the assessment process:

I determine what and how to assess, I don't let the students determine it.... also I do not think we should let kids decide how to assign grades; it won't be real.... everyone will take advantage of it, and I do not think that's right. We [the teachers] see the broad picture, the global one, and they don't. This is my opinion. I've never tried, but I don't think it's right.  
(Teacher, School T1)

By contrast, the holders of an AC mindset believe that assessment should empower the learner, especially through feedback by the teacher or peers:

What stands before my eyes [when I write feedback notes] is the student, I put myself on the other side [and think] what I would want, how I would want others to talk to me when I get back my graded work; which accents, reinforcements... and most importantly, how to empower the student, to give him the confidence to face the next task.

(Teacher, School A3)

Student (P): first you [student giving peer feedback] write the good comments and then the not so good ones...

Interviewer: Why do you write the good comments first?

Student (B): This warms his heart... after he knows he is doing good things he... knows that he also needs to learn things.

(Second graders, School A1)

**Attitudes toward diversity.** The holders of a TC mindset believe that in assessment 'one size fits all' and that assessment ought to follow a strict plan and avoid circumstances that give rise to uncertainty:

I have a rigorous assessment program, it consists of three cycles: immediate knowledge, which is examined in a 'success test' and shows the extent of absorption of the immediate material (short-term memory); knowledge at the topic level that is examined in the 'control test'; and knowledge acquired over time that is examined in the 'final test.'

(Teacher, School T2)

By contrast, the holders of an AC mindset acknowledge student diversity and assess informally on a regular basis:

I assess students regularly. In informal verbal dialogue...in one-on-one conversations with students. Assessment is done constantly, until it reaches the stage where it needs to be expressed in writing.

(Teacher, School A3)

Moreover, they have a positive attitude towards uncertainty:

you know where you begin the lesson, but you never know where you finish it, because there is a strong need to respond according to factors that arise during the lesson. So even if I'm well prepared for the lesson I do not always know where it will go.... For me personally, this is what fascinates me...

(Teacher, School A2)

**Expectations about learning.** The holders of a TC mindset believe that ability is fixed (innate); hence, they have low expectations for low-achieving students and consequently provide low-quality instruction and assessment practices:

The reality of weak students restricts my teaching options, such as conducting mutual feedback from student to student [peer assessment]. Their capabilities are low and that is where I give up...  
(Teacher, School T4)

By comparison, the holders of an AC mindset express efficacy beliefs; teachers believe in their students' ability to learn and in their own self- and collective ability to advance the learning of all of their students. Likewise, the students believe in their ability to learn. The following quote attests to the principal's role in enhancing both teachers' and students' efficacy beliefs:

Most salient [about our school principal], in my opinion, is her belief in the ability inherent in every person, a student and a teacher. She is not willing to accept that a student cannot [learn], she has embedded in the staff the belief that if a student fails it is because we have not found the way to reach him/her, and should keep looking. This is what is happening today in the school; mapping, monitoring each student's [progress], attempts to support ... and lead him/her in creative ways; if unsuccessful in one way, we'll try another [way]... it's something that came from her and has been internalized by teachers and students alike.  
(Teacher, School A3)

**The fidelity of assessment.** The holders of a TC mindset believe that assessment is accurate quantification: 'I believe in tests, in my opinion, a test is the most dependable tool to measure students' knowledge' (Teacher, School T1).

In contrast, the holders of an AC mindset believe that tests can be faulty sources of information and that modesty in assessment is therefore required:

A test is not necessarily a trustworthy indication about student knowledge because maybe the student did not understand the wording of the questions, the instructions were not clear for him, or he was not feeling well on the day of the test. Therefore, you have to be cautious when interpreting test scores.  
(Teacher, School A2)

## ***16.3.2 Classroom Assessment in AC and TC Schools***

### **16.3.2.1 Formal Assessment**

AC schools are student centered and committed to the learning of all of the students. As stated by an elementary school principal, 'It is highly important for me that teachers be student oriented and have responsibility for the learning of all students' (School A2). As such, AC schools, compared to TC schools, place a greater emphasis on formative assessment and exhibit AfL praxis of relatively high quality with regard to each stage in the AfL cycle: planning; developing an assessment tool; collecting evidence; analyzing and interpreting student

performance; acting to close the gap between expected and obtained performance; and evaluating the effectiveness of the actions. AC schools are attentive to the alignment among AfL stages and the regularity of applying the cycles. Moreover, the measures that are taken to foster self-regulated learning are evident in assessment-related artifacts, such as teachers' feedback notes, which include scaffolding in the form of questions or clues to help students correct their answers; rubrics for self-assessment; and students' reflective notes.

Students take an active part in the assessment process, as described by a third-grade teacher:

Regarding rubrics:

We take samples [of student writing], analyze them, then, together we see what is better and what is worse, we explain why, they [the students] explain—not I—and then I say: 'okay, from all that we've done, let's now say what is needed for it [the written piece] to be good.' It comes from them.... At the most, if they forgot a point or two, I'll add it...

Regarding peer assessment:

I let kids assess each other's work. We prepare a rubric... we see exactly what we are going to assess and then in pairs they check each other's work according to the rubric... and write what is right and mark what needs to be fixed, what needs improvement; and then they do a second draft based on what the peer wrote to them and give it to another peer or the same one to check... until the product is ready... the kids really like the process.... Eventually I show them the texts they wrote at the beginning and the texts they wrote at the end, we analyze and see the difference...

(Teacher, School A1)

By contrast, in TC schools, less coherence of practice is noted, ranging from misinterpreting AfL as frequent summative assessment whereby items from standardized tests are administered within short time intervals and used to report students' achievements, to superficial implementation of certain AfL strategies within a quantitative assessment framework. This condition is illustrated in the following description of a sixth-grade teacher talking about peer assessment of portfolios in her class:

[A]ll their [students'] products are filed in portfolios.... I grade the portfolios by the number of items they contain, actually the students do it, they do peer assessment.... towards the end of the term I open the cabinet where the portfolios are stored... they [students] take a sheet of paper and write down what should be in each divider in the portfolio.... I do not check it [the list]. I divide them into pairs without them knowing in advance who [will work] with whom... and they check each other's portfolio. They have the list, if the item exists they mark ✓ if it is missing they mark X. Then, they divide 100 by the number of items in the list and multiply by the number of ✓'s they marked so they really give each other a grade. And I use it for the report card.... They are very excited about this [activity] .... They love giving each other grades...they really enjoy it.

(Teacher, School T1)

However, when this teacher explains why she does not involve her students more in peer assessment, she claims: 'frankly, I don't believe in it, it does not seem logical, they will obviously give 100 to the friends that they like and will harshly treat those that they don't like.... They are too young for this... they lack professional knowledge. Even I find it difficult to accurately assess student

performance, let alone them...’ Consequently, she does not involve her students in self-assessment, nor in any constructive decisions regarding assessment.

Generally, for graded tests in these schools, a total score is written on top of the page, above the teacher’s signature, and in the margins, the correct answers are marked by ✓’s and incorrect ones, either by ✗’s or by the number of points deducted.

### 16.3.2.2 Classroom Culture

In AC schools, the dominant attributes of classroom culture reflect constructivist conceptions of instruction, learning, and assessment: a sense of student agency, a learning orientation, emphasis on higher-order thinking skills, dialogue, reflection, norms of collaboration, transparency, tolerance of errors, and recognition of their value to advance learning. Mutual trust and respect, caring, openness, and pride characterize the interpersonal relations between teacher and students and among students in AC schools.

More variability in classroom culture is evident in TC schools, where the climate ranges from toxic and highly competitive, reflecting an intolerance of errors (e.g., teasing students who flunked a test) to a climate that is characterized by a lack of enthusiasm, feelings of alienation from teachers, coupled with students’ obedience and willingness to do what is required to get good grades.

The following four excerpts from interviews with focus groups of students in the AC and the TC schools (excerpts 1, 3 and 2, 4, respectively) exemplify some of the above-mentioned characteristics:

#### *Excerpt 1.*

Interviewer: How did you feel when R [the teacher] projected a slide of what you wrote?

Student (B): I felt fine... I know that even if I get bad comments they are just so that I could make more progress in my studies.

Interviewer:... Do you think it’s important to project someone else’s writing and respond to it?

Student (T): I think it is important, because then you can learn a lot of things for your next descriptions.

....

Student (A): She’s right, because from all sorts of descriptions R is projecting, with mistakes and without mistakes, with beautiful descriptions and not so successful [descriptions], we learn things that are important.

(Second graders, School A1)

#### *Excerpt 2.*

Student (A): There are kids who receive a low grade and they do not want to show other kids, they hide it...

Interviewer: Why, why would a kid want to hide the grade?

Student (A): Because it is a low grade.

Interviewer: Why does he want to hide it?

Student (E): Because he is ashamed.

Student (A): So that they [other students] do not laugh at him.

...

Student (E): Because they [other students] want to grab his test.

Interviewer: Why do they want to grab it?

Student (E): Because they want to disclose it.

Student (A): To the entire class.

Interviewer: Why?

Student (E): So everyone knows that he is not smart...

(Second graders, School T2)

*Excerpt 3.*

Student (M): She [the teacher] really cares for us and will do everything so we can succeed.

Student (N): Any time you can call her and she always helps.

Students (A): It's not that the school day is over and nobody knows us or forgets us.

(Ninth graders, School A3)

*Excerpt 4.*

Student (N): The teacher comes to class, starts teaching the material, then [assigns] homework, and goes.... The teacher does not say good morning to us.... I'm not there for her. Teachers do not care about the students. I feel like she does not see us, we are air.

(Ninth grader, School T4)

### **16.3.3 Teacher Professional Learning (TPL) in AC and TC Schools**

In AC schools, TPL is focused on student learning to promote their agency to self-regulate learning. To that end, the staff is continuously engaged in collaborative inquiry into their practice in order to improve it.

Efficient mechanisms and regularities to facilitate TPL are offered: professional meetings are held regularly; teachers observe the lessons of their peers and provide feedback; teachers consult with their peers when they face pedagogy-related difficulties or invite them to their lessons to demonstrate a technique in which they are proficient. Such actions attest to norms of transparency, the 'deprivatization' of practice, and legitimization of errors. Furthermore, all of the teachers are members of one or more teams; some professional collaborations are set out in the school's annual working plan, others are ad hoc. In addition, some of the staff members attend external PD programs that are aligned with the school's needs and then lead the learning of those topics in their teams.

A positive social climate is evident in team meetings, as well as deep discussions. The interviews with staff members reflect shared responsibility coupled with high expectations for the learning of all students in the school, and professional self- and collective efficacy. References are made to a shared school vision, and a common language is manifested in the discourse.

TPL emerges from the specific needs of the school and addresses the areas that have generated discomfort among the staff. Full inquiry cycles are conducted, for instance, with regard to the ideas or concepts that students have difficulties comprehending. In AC schools, collaborative inquiry processes follow a variation of



'Lesson Study' (Fernandez 2002), whereby a team of teachers who teach the subject designs a task to elicit students' understanding. Based on an analysis of students' performance on the task, the team develops a re-engagement lesson to address the identified misunderstandings and misconceptions. A team member then teaches the re-engagement lesson while the others observe it, focusing on students' understanding. The effectiveness of the lesson is then evaluated using a post-treatment task.

By contrast, in TC schools TPL assumes only a marginal role. Moreover, in these schools, a larger variability in external frameworks for TPL is evident. Teachers attend the external PD programs of their choice or are told by the principal which external PD program to attend; alternatively, PD takes place in school by external providers who are selected by the principal. Opportunities to build on local expertise when offered are often rejected, as one elementary teacher complains:

Teacher: Someone has to understand that in our school there are good forces.... that should be given the opportunity to lead PD in an area they are good at...

Interviewer: What needs to be done to make this happen?

Teacher: ... I know that today we bring an [external] instructor to teach us. When she comes suddenly there is time, ample time, clearing all teachers from their classes to sit and listen to her advice... I sit there and freak out; I studied the method... for my M.A. degree and have already implemented it in my class...

Interviewer: So why is this happening?

Teacher: Because there are no hours. I have to be in class, that's what I get paid for. But when an external instructor comes, suddenly there are hours. I see what she teaches and it infuriates me. I could have done it better in our team ...

(Teacher, School T1)

As to pedagogical staff meetings, evidence from internal memos indicates that although such meetings are scheduled at the beginning of the school year, quite a few are cancelled, as indicated by a middle school teacher:

I think the team should meet regularly once a week. Regularly is not something that takes place sometimes during breaks or somehow like that, just to mark a ✓. It's true that we have an organized schedule for staff meetings... but most of the time they are cancelled because of conflicting events... but I think there is also a problem of teachers' willingness to be flexible and make an effort to come to the meetings.

(Teacher, School T4)

When pedagogical meetings do take place they are devoted mainly to exchanging information and making decisions regarding technical or administrative matters such as grouping students, reporting grades, compiling comparative tests, and choosing textbooks.

Observational evidence from staff meetings indicates variability in the social climate, ranging from depressing (dominated by accusations, complaints, and top-down instructions) to pleasant and friendly (more like a social gathering than a working meeting).

### ***16.3.4 Leadership in AC and TC Schools: Effects on TPL***

In AC schools, individual and collective capacity building is considered to be a top priority by the principal who considers him- or herself to be a pedagogical leader, hence, many efforts and resources are invested in facilitating school-based PD. As contended by a school principal (School 4A), ‘The teachers are “my classroom,” and my efforts are targeted toward their professional development.’

Affordances to facilitate TPL include learning mechanisms and regularities for staff participation, for monitoring and evaluating, and for information management. The principal not only facilitates TPL but also takes an active part in it.

Moreover, the organizational structure allows every teacher to assume an active leadership role. This distributed leadership is highly encouraged by the principal and supported by existing norms such as the legitimization of errors. As a middle school teacher explains:

Our school motto is that if you try, you may make mistakes; yet learning from mistakes facilitates growth. I think that when the school principal sends such a message to the whole staff, encouraging them to initiate, try, make mistakes, and ponder, it becomes common practice. Today I feel that this is something that the staff believes in, and passes it on to our students...

(Teacher, School, A3)

In addition, ad hoc collaborations are formed, thus increasing the learning interactions among staff members.

An inquiry disposition and constant reflection propels organizational learning in these schools. As a school principal attests:

We are constantly checking the achievement of our goal. Did we reach it? Did we not reach it? What needs to be done? What does not need to be done? We are constantly in this business. This, by the way, this is a school culture; we are not waiting for the school year to end. Our reflective story says: We did this and that, what results did we get? Do we feel good about it? Do we feel bad about it? Should we change it? Should we be doing something else?

(Principal, School A4)

By contrast, in TC schools, a more centralized structure and autocratic leadership are evident; teachers have fewer opportunities to participate in decision making and to assume leadership roles. The management tends to be outcome oriented and invests efforts in measuring, ranking, and comparing students’ grades and seems to be less interested in exploring what caused them, thus, it refrains from investing effort in facilitating school-based PD.

Consequently, excessive testing is performed, and teachers spend much time grading tests and entering scores into a central information system that enables the management to review the results at any given moment and to generate reports to stakeholders (i.e., parents, superintendent, and local authority). No time and effort are devoted to the teachers’ collaborative analysis of students’ performance to diagnose their misunderstandings and misconceptions and to use the information to design remediation interventions.

### ***16.3.5 The Impact of External Accountability Tests in AC and TC Schools***

The external accountability tests that are administered in elementary and middle schools (grades five and eight) in Israel are called the GEMS (Growth and Effectiveness Measures of Schools) tests. The evidence that was collected in the AC and TC schools about the impact of these tests with regard to preparation for the tests and the utilization of their results indicates several differences. TC schools tend to be more focused on preparing for these tests while utilizing their results less than is the case in AC schools. In TC schools, the curriculum is altered in classes that are scheduled to take the tests to reinforce the tested subjects at the expense of other school subjects, and the students are immersed in intensive drill-and-practice on items that are similar to those that have appeared on GEMS tests in previous years. In contrast, in AC schools, no special test preparation arrangements are made. Instead, the staff studies the results of the previous GEMS administration upon receipt, and conclusions are drawn and incorporated systemically in the school curriculum to reinforce the topics in which performance on the GEMS test was unsatisfactory. In contrast, when the reports of the GEMS results arrive at TC schools, the staff is only informed, and its members are praised or rebuked accordingly, yet no systemic inferences are drawn to improve practice.

Differences are also identified with regard to climate in the classes that are scheduled to take the GEMS tests. In TC schools, students complain of excessive homework and practice tests, whereas teachers complain of extreme pressure from the school management to cover the parts of the curriculum that are addressed in the tests. They also complain of the stressful conditions under which they operate—on their own, with no support from other staff members, realizing that they alone will be held responsible for the results. By contrast, in AC schools, the GEMS tests have no noticeable effect on the classroom climate as no special preparatory measures are taken for these tests.

Differences are also identified with regard to reactions to the GEMS results. In TC schools, excuses are primarily heard, blaming unsatisfactory results on factors that are beyond the schools' control, such as the students' abilities and backgrounds. By contrast, AC schools' responses mainly take the form of actions to reassess the students in areas in which their performance on the GEMS tests is unsatisfactory to diagnose misunderstandings and to provide remedial instruction.

## **16.4 Discussion**

As shown in the results section, the comparison between the AC and TC schools reveals clear differences with regard to culture attributes (conceptions, beliefs, expectations, orientation, values, norms, and social climate) which are evident at the classroom and the school levels and reflect distinct AC and TC mindsets.

Differences between AC and TC schools are also noted in organizational structure, leadership practice, TPL, and classroom assessment practice, as well as in patterns of response to external accountability demands.

AC schools are complex, learning-oriented, and student-centered systems, whose aim is to improve teacher practice and student learning. Their organizational structure allows teachers to participate in pedagogy-related decision making and to take an active leadership role. The principal considers investment in capacity building to be a top priority and thus facilitates TPL by establishing supportive learning mechanisms and takes an active part in the learning process. Teachers meet regularly to inquire collaboratively into their practice, analyzing evidence from students' written work, from lesson-documented observations as well as from surveys and interviews with students. Collective high expectations for all of the students shape the teachers' work, which is characterized by pedagogic consistency. Teachers set goals that are derived from a shared vision and that are adapted to the specific needs of their student population. They monitor their actions toward attaining the goals by applying their shared understanding of good practice. They exhibit resilience in view of external demands and are not distracted by high-stakes external assessment, hence demonstrating professional accountability (O'Day 2002).

Learning in the classroom and in professional staff meetings occurs through participation and is driven by assessment that is aimed at improvement. A spiral assessment process is carried out and links AfL to inquiry cycles (Birenbaum et al. 2009). Assessment is continuous, complex, contextual, and its parts are interconnected. An inquiry disposition is evident among the staff members as they inquire into their practice and among students who inquire into their performance, asking questions to improve upon it, such as, 'Why did I get this answer wrong?', 'What am I doing that is incorrect?' No wonder they prefer feedback to grades. As a second grader explains:

When you get comments [feedback] it's better... because when you get comments you understand your mistakes and can correct them... comments are just for your own good... so no need to be offended, you only learn from them.  
(Second grader, School A1)

A similarity between CL and TPL is also noted with respect to empowerment, efficacy beliefs (Bandura 1997), and causal attribution (Weiner 1992). The teachers are empowered by the principal, and the students by their teachers. High self- and collective efficacy beliefs are expressed by the teachers regarding their ability to teach low achievers and by students regarding their ability to learn. The staff and the students declare that they know 'where we are, where we are going, and how to get there.' The teachers and the students alike attribute their success or failure to causes that are within their control.

By contrast, TC schools are principal-centered bureaucratic organizations (Sergiovanni 1995) in which teachers work mostly in isolation and are subject to top-down dictates. This principle also characterizes the power relations between the teachers and the students, who are submissive subjects to instruction and

assessment practices. The schools are test driven and subject-matter centered, concentrating on quantitative comparisons of outcomes rather than trying to understand what caused them. Questions like ‘How many answers did I get wrong?’ or ‘Have the test scores increased or decreased compared to last year?’ characterize the assessment discourse in these schools. It is not surprising that their students prefer grades to feedback, as expressed by sixth graders:

Interviewer: If you were told, ‘beginning tomorrow, no grades... you will get only feedback,’ what would you say?

Student (M): A grade is important to me, I want to know a number; I want the grade as a number: 94, 82.

Student (B): Grades are indispensable. Grades come to you directly...

Student (S): I prefer the grade itself instead of feedback.

Student (K): It is important to know exactly where we stand.

(Sixth graders, School T1)

School-based TPL is not promoted, relying instead on teacher enrollment in external PD programs. Similarly, external instructional aid is often employed to provide remediation for underachievers. Linearity characterizes teacher practice in these schools (e.g., teach—examine—produce a grade—report it). Assessment processes are simplistic, disconnected, decontextualized, and often missing their formative function. Overall, TC schools are characterized by fragmentation and a weak agreement around values, norms, expectations, and processes. They lack a systemic approach to instruction and assessment and a shared understanding of good practice, which results in variations in practice among teachers, as evidenced in the following quote:

my goal is not only the grade... it is important to me that they [the students] will understand; it is important to me to develop something more.... Teachers who do not see themselves in this place, I cannot convince her to do portfolios, because she [her counterpart] does not understand. She understands: I came, I taught, I gave a test, there is a grade; I need to know his grade, he [the student] needs to know his grade, his parents need to know, and that’s all. So that’s why she chooses to take this course of action.

(Teacher, School T1)

TC schools adhere to bureaucratic, outcomes-based accountability (O’Day 2002). They follow externally set goals without attempting to adapt them to their specific needs. Acting as submissive subjects to external dictates tends to cause stress and despair, as attested by an elementary school teacher: ‘The management is under a lot of stress.... The superintendent places incredible stress on us. It starts where she [the superintendent] stresses the principal, who stresses us...’ (School T2). Finding that TC schools are negatively affected by high-stakes external assessment is thus not surprising.

The TC schools show a predominance of positivist epistemology and traditional conceptions of instruction, learning, and assessment, as well as expressions of external motivation and intolerance of errors. Low self-efficacy beliefs are often expressed by the teachers with regard to their ability to teach low achievers and by students with regard to their ability to learn. Both teachers and students tend to attribute their success or failure to causes that are beyond their control.

Taken together, our findings point to an underlying dimension of internal coherence along which AC and TC schools differ. Elmore et al. (2014) define internal coherence (IC) as ‘a school’s capacity to engage in deliberate improvements in instructional practice and student learning across classrooms over time, as evidenced by educator practices and organizational processes that connect and align work across the organization’ (p. 3). Their IC model for school improvement addresses three interrelated and reciprocally reinforcing dimensions: leadership practice for instructional improvement; organizational processes (at whole-school and team levels), and teacher efficacy beliefs (individual and collective). In light of the definition and conceptual framework that was advanced by Elmore et al. (2014), it can be inferred, according to our evidence, that AC compared to TC schools demonstrate a higher degree of internal coherence.

From a complexity framework, what seems to enhance internal coherence in AC schools are the continuous reciprocal interactions within and between the three systems that are nested in school culture (CL and TPL, as well as school leadership). According to complexity principles, the more frequent and powerful the interactions among agents, the more influence they are likely to have on agents’ behavior and consequently on the emergence of new patterns, hence, the more likely the organization will be to experience renewal (Mitleton-Kelly 2003). To illustrate this principle, TPL is often triggered by difficulties that are encountered in classrooms; team discussions yield solutions, which are then applied in classrooms. The experience that is gained from the implementation is shared within the team, which critically reviews it, draws conclusions, and suggests adjustments and refinements accordingly. The improved version is then implemented in classrooms, and the process repeats until a satisfactory outcome is reached, thereby enriching the individual and collective repertoire of practices. By empowering emergent individual expertise, the reciprocal interactions between TPL and CL expand to the leadership system, making distributed leadership more salient and effective. Moreover, the continuous reciprocal interactions also seem to affect teacher efficacy beliefs, both individual and collective. The four sources of efficacy-shaping information that were postulated by Bandura (1997), namely, mastery experience, vicarious experience, social persuasion, and affective state, emerge through these interactions. Moreover, the empowering interactions also seem to affect teachers’ causal attributions (Weiner 1992) regarding student performance (i.e., whether causes are within or beyond their control). It should be noted that according to Bandura (1997), efficacy beliefs and causal attributions are related in a reciprocal manner.

In view of the above, it is quite obvious why AC, as opposed to TC, is conducive for a proper implementation of AfL. However, schools where AC dominates are a minority in Israel. In most schools, TC dominates or both cultures (AC and TC) coexist, either because the school culture includes elements of both or because there are subgroups of teachers who adhere to one or to the other. In such cases, the question becomes how to develop the place of AC and help all teachers contribute to a school culture that fosters student learning.

## 16.5 Implications for Action and Future Directions of Research

The main implication of the study is that viewing AfL through a wide-angle lens of school culture enables us to see what is required to fulfill AfL's potential to promote student learning. Moreover, it shows that the potential of AfL to inform (and transform) practice transcends CL to encompass school-based TPL, and highlights the critical role of school leadership in facilitating it.

### 16.5.1 Recommendations for Developing AC

The following recommendations are based on two premises: (a) AC is a complex system in which three other systems (CL, TPL, and leadership) are nested; therefore, an improvement plan should address all four systems; (b) schools vary along the cultural continuum; therefore, an improvement plan of the 'one size fits all' type is unrealistic; instead, a plan tailored to a school's needs is required. Consequently, schools that aim to develop AC should first assess their standing with regard to culture, TPL, CL, and leadership.

As to the improvement plan, in schools that are located closer to the TC end of the culture continuum, attention should be given first to the principals, offering them PD to help them understand their role in developing AC. It is recommended that they meet with principals from AC schools and if possible also be mentored by them.

PD in AfL for teachers should be contextualized and tailored to a staff's level of functioning as a school-based professional learning community. Schools with a relatively high level of functioning, but which are not yet familiar with AfL, would be advised to conduct a variation of a Lesson Study (Fernandez 2002), as described in Sect. 16.3.3 above. Such a process would provide 'hands-on' experience in formal and informal AfL, which should then be supplemented with a formal introduction to AfL (addressing its rationale, principles, and practices, as well as the conditions that facilitate it and those that hinder it). Teachers' attention should be drawn to the similarity between AfL cycles that are applied at the classroom level and inquiry cycles that are conducted at the whole-school and team levels.

For schools with a low level of functioning as a professional learning community, PD should first focus on teamwork and introduce the staff to its benefits, preferably by team members from AC schools. Hands-on in AfL would be gradually introduced, starting, for instance, with a collaborative analysis of students' test performance, working to diagnose the sources of incorrect responses and providing feedback to help students understand their mistakes. Introducing teachers to instructional tools that help focus their attention during the lesson on student understanding (such as traffic light cards, by which students signal their level of understanding; erasable mini boards or ICT tools, by which all of the students'

responses to teachers' questions are shown) are additional means to shift teachers' focus from testing for grading to assessing for learning. Next, the teams would be introduced to AfL and advised to collaboratively engage in designing performance tasks including rubrics. They would then analyze a sample of students' performance, practice writing feedback notes, and make inferences regarding the attainment of instructional goals. Further on, the teams would be advised to conduct an inquiry cycle of the form that is suggested above (Lesson Study), which would also provide hands-on experience in informal AfL.

### 16.5.2 Recommendation for Further Research

It is recommended to design interventions to assist schools in developing AC and to investigate the transformation process that schools go through and the challenges that they face as they build staff capacity to promote student agency. It is also recommended to conduct international comparisons of assessment-related school cultures, with respect to CL, TPL, and leadership to discern the multilevel contextual effects on teacher assessment practice, specifically as it promotes learning.

**Acknowledgments** I would like to thank Ayelet Keren, Helena Kimron, Yamit Markovich, Simon Sadres-Leibovich, Neria Shahor, and Hany Shilton for providing the data on which this paper is based.

## References

- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman and Company.
- Birenbaum, M. (1996). Assessment 2000: Towards a pluralistic approach to assessment. In M. Birenbaum & F. J. R. C. Dochy (Eds.), *Alternatives in assessment of achievements, learning processes and prior knowledge* (pp. 4–29). Boston, MA: Kluwer Academic Publishers.
- Birenbaum, M. (2014). Conceptualizing assessment culture in school. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *The enabling power of assessment: Vol. 1. Designing assessment for quality learning* (Vol. 1, pp. 285–302). Dordrecht: Springer.
- Birenbaum, M., Kimron, H., Shilton, H., & Shahaf-Barzilay, R. (2009). Cycles of inquiry: Formative assessment in service of learning in classrooms and in school-based professional communities. *Studies in Educational Evaluation*, 35, 130–149.
- Black, P. J., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, 5(1), 7–74.
- Dweck, C. S. (2006). *Mindsets: The new psychology of success*. New York: Random House.
- Elmore, R. F., Forman, M. L., Elizabeth, L., Stosich, E. L., & Bocala, C. (2014). *The internal coherence assessment protocol & developmental framework: Building the organizational capacity for instructional improvement in schools*. Washington, DC: Strategic Education Research Partnership (SERP).
- Fernandez, C. (2002). Learning from Japanese approaches to professional development: The case of lesson study. *Journal of Teacher Education*, 53, 393–405.



- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage.
- Marshall, B., & Drummond, M. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education*, 21(2), 133–149.
- Mitleton-Kelly, E. (2003). Ten principles of complexity and enabling infrastructures. In E. Mitleton-Kelly (Ed.), *Complex systems and evolutionary perspectives of organizations: The application of complexity theory to organizations* (pp. 23–50). Oxford: Elsevier.
- Murphy, M. C., & Dweck, C. S. (2010). A culture of genius: How an organization's lay theory shapes people's cognition, affect, and behavior. *Personality and Social Psychology Bulletin*, 36(3), 283–296.
- O'Day, J. A. (2002). Complexity, accountability, and school improvement. *Harvard Education Review*, 72(3), 293–327.
- Peterson, K. D., & Deal, T. E. (1998). How leaders influence the culture of schools. *Educational Leadership*, 56(1), 28–30.
- Sergiovanni, T. J. (1995). *The principalship: A reflective practice perspective* (3rd ed.). Boston, MA: Allyn & Bacon.
- Weiner, B. (1992). *Human motivation: Metaphors, theories and research*. Newbury Park, CA: Sage.

# Chapter 17

## The Role of Classroom Assessment in Supporting Self-Regulated Learning

Heidi Andrade and Susan M. Brookhart

**Abstract** Self-regulation of learning occurs when learners set goals and then systematically carry out cognitive, affective, and behavioral practices and procedures that move them closer to those goals. Self-regulated learning (SRL) depends, in part, on information gleaned from classroom assessments about student learning and achievement. In this chapter we will discuss how classroom assessment is or could be used to support SRL. We will draw on the literatures on classroom assessment and SRL in order to demonstrate how assessment contributes to each phase of self-regulation, defined here as: (1) goal setting, (2) progress monitoring, and (3) revision and adjustment. For example, the goal-setting phase is influenced by the learning goals and success criteria shared by a teacher. The progress-monitoring phase is affected by feedback provided via formative and summative assessments. The revision-and-adjustment phase is affected by opportunities teachers give students to use feedback and decisions students make based on that feedback. This chapter demonstrates the close relationship between classroom assessment and SRL by reviewing research evidence for each phase, and makes the case that assessment can support the self-regulation of learning in classroom settings. The chapter also addresses challenges of implementing classroom assessment practices that support SRL.

### 17.1 Introduction

Classroom assessment includes assessment by classroom teachers for both formative and summative classroom purposes. Formative classroom assessment strategies, or assessment for learning (AfL), are used by both teachers and students to

---

H. Andrade (✉)

University at Albany, State University of New York, Albany, USA

e-mail: handrade@albany.edu

S.M. Brookhart

Brookhart Enterprises, Helena, USA

S.M. Brookhart

Duquesne University, Pittsburgh, USA

© Springer International Publishing Switzerland 2016

D. Laveault and L. Allal (eds.), *Assessment for Learning: Meeting the Challenge of Implementation*, The Enabling Power of Assessment 4,

DOI 10.1007/978-3-319-39211-0\_17

further student learning. Summative classroom assessment, or grading, is usually a function reserved for the teacher and is done for the purpose of certifying and reporting learning. These functions blur a bit in the classroom context, as students use both kinds of assessment to inform the self-regulation of learning. Thus classroom assessment is one of the other sources, besides the self, that influence the regulation of learning (Allal 2010). This chapter examines how features of classroom assessment such as success criteria, feedback, and opportunities for revision influence the process of self-regulation of learning and, therefore, influence student learning.

## 17.2 Self-Regulation of Learning

Self-regulation of learning occurs when learners set goals and then systematically carry out cognitive, affective, and behavioral practices and procedures that move them closer to those goals (Zimmerman and Schunk 2011). Current theories of formative assessment also recognize that the agency for learning resides with the student (Andrade 2010). Self-regulated learning (SRL) depends, in part, on information gleaned from classroom assessments about student learning and achievement: This is a likely explanation for research findings that suggest formative assessment is a potential influence on SRL (Nicol and Macfarlane-Dick 2006). In this chapter, we discuss how classroom assessment in general and formative assessment in particular are, or could be, used to support SRL.

Scholarship on self-regulation organizes cognitive, metacognitive, and motivational aspects into a general view of how learners understand and then pursue learning goals. Different theorists have presented models of how students activate cognition, metacognition, and motivation in order to learn. Three influential models are a nested view (Boekaerts 1999), an information processing view (Winne and Hadwin 1998), and a phase or cyclical view (Pintrich and Zusho 2002; Zimmerman 2011). Phase views of SRL allow theorists to place cognitive, metacognitive, and motivational constructs into a sequence of events that occur as students self-regulate. For example, Pintrich and Zusho (2002) organize the phases and areas of self-regulation into four phases: (1) forethought, planning, and activation, (2) monitoring, (3) control, and (4) reaction and reflection, which include making attributions of success or failure. For another example, Winne (2011) describes self-regulated learning as ‘unfolding over four weakly sequenced and recursive phases’ (p. 20). In Phase One, the learner defines the task and its affordances and constraints. In Phase Two, the learner sets goals and plans. During Phase Three, the learner engages with the task, and in Phase Four the learner evaluates his or her work, which can result in making revisions or adjustments to the work.

A phase view of SRL affords a way to crosswalk the classroom assessment literature, since assessment also can be described as having three main phases, also

cyclical in nature: (1) goal setting, (2) monitoring via feedback, and (3) revision or adjustment. The similarities between the phases of SRL and classroom assessment are clear: Both SRL and classroom assessment involve setting goals, monitoring/evaluating progress toward those goals, and reacting to feedback about gaps between goals and progress by making adjustments to teaching, learning, and/or work products. As with SRL, the differences between the monitoring and revision/control phases make sense conceptually but are difficult to separate empirically.

SRL and classroom assessment, especially formative assessment, have overlapping aims but distinct bodies of research and classroom practices—at least until now. A goal of this chapter is to better understand how what we know about SRL can inform classroom practice, and vice versa. We will do so by examining research on the relationship between SRL and the three phases of classroom assessment: (1) goal setting, (2) progress monitoring, and (3) revision and adjustment. There are currently only a few studies that directly examine this relationship, but the results are promising.

### **17.3 Classroom Assessment and Self-Regulated Learning**

Successful formative assessment practices work because they support learner autonomy (Black et al. 2006). Feedback gleaned from classroom assessments can provoke students to self-regulate their learning by providing evaluations of their understanding and performances. Ideally, those evaluations lead students to assess whether particular strategies are effective in meeting their learning goals and to make adjustments to their knowledge, motivation, behavior, and even context. Under the right conditions, sources of feedback include not only teachers but also students themselves, their peers, and computer-based technologies designed to deliver instant automated feedback. In the remainder of this chapter, we describe how feedback from a variety of sources can or could be used to support self-regulated learning. We show that successful students use formative assessment information, as well as some information from classroom-based summative assessment, in support of the self-regulation of learning.

Our focus is on cognitive processes but it is important to note that motivation is also an important component of SRL and a phenomenon that is highly susceptible to influence from assessments, particularly summative grades. Unlike formative feedback, summative assessment has gained a reputation for having unintended, often destructive consequences for both learning and motivation. For example, research showing that grades negatively influence performance and motivation (Butler 1987; Butler and Nisan 1986; Lipnevich and Smith 2008) implies that grades can trigger counterproductive regulatory processes, especially for low-achieving students.

## 17.4 Phase One: Goal Setting

Classroom assessment serves a clear purpose in terms of goal setting. Perhaps the most obvious instantiation is the setting of learning goals by teachers. Various called learning intentions, learning goals, and learning targets in current scholarship on assessment and self-regulated learning, goals describe the skills, concepts, and dispositions that constitute the intended consequences of teaching and learning. Modern theories of regulated learning consider goals to be fundamental to regulatory proficiency and success (Hadwin et al. 2011; Winne 2011; Zimmerman 2011), and theories of classroom assessment consider teachers' learning goals for students to be the basis of good assessment (Allal 2010; McMillan 2011; Nicol and Macfarlane-Dick 2006; Stiggins 2008). If feedback is to be beneficial, students must have a clear understanding of the goal or standard for a performance, and be able to compare their performance with that standard, after which they can take relevant action in order to close any gaps (Hattie and Timperley 2007).

Portfolios are a form of classroom assessment that tends to highlight goal setting by students. Although quite limited, the research on portfolios suggests a positive relationship between goal setting and students' performance (Belgrad 2013). Ideally, teachers and students discuss the goals to be attained, as well as the criteria and standards for particular assessments (Allal 2010).

The classroom assessment literature places a special emphasis on success criteria (Heritage 2010; Moss et al. 2013; Torrance and Pryor 2001). In contrast with learning goals, which tend to be broad, success criteria describe the qualities of excellent student work on a particular assignment. Success criteria can be communicated to students in a variety of ways, including via rubrics (Andrade 2000; Brookhart 2013a), exemplars, or worked examples that imply success criteria (Hattie 2009). Success criteria can be co-constructed with students, as for example when teachers and students together create a rubric (Andrade et al. 2008).

In a study that employed rubrics and/or exemplar research proposals, Lipnevich et al. (2014) found that providing 100 undergraduates<sup>1</sup> with rubrics, exemplars, or both was associated with significant improvements in student performance, with rubrics edging out the exemplars only and rubrics + exemplars conditions in terms of effect size (rubrics only Cohen's  $d = 1.54$ ; exemplars only Cohen's  $d = 1.04$ ; rubrics + exemplars Cohen's  $d = 1.04$ ). Similarly, Andrade and her colleagues (Andrade et al. 2008, 2010) found that providing elementary and middle school students with model papers and rubrics, combined with a scaffolded process of self-assessment, was related to statistically significant and practically meaningful differences between the performance of the students in the treatment and comparison groups (elementary school Cohen's  $d = 0.87$ ; middle school Cohen's  $d = 0.66$ ). Taken together, these studies strongly suggest that sharing success

---

<sup>1</sup>Throughout the chapter, the grades mentioned correspond to the U.S. K–12 system (or equivalent systems elsewhere), with students 5–18 years old, and the term undergraduates refers to students in Bachelor level university studies.

criteria with students can promote learning in both primary school and higher education.

Students' understandings of their teachers' criteria can influence their regulation of their learning (Butler and Cartier 2004; Butler and Schnellert 2015). For example, a student who interprets a reading assignment as a memorization task will plan to use low-level cognitive processes and consider himself or herself successful once key terms are memorized. In contrast, another student who interprets the same assignment as requiring comprehension and application will employ sophisticated comprehension strategies and define success as being able to explain and transfer the content of the text. Thus, teachers cannot assume that success criteria will be uniformly adopted and applied by students: Efforts must be made to ensure accurate and effective interpretations of the criteria (Butler and Cartier 2004).

Students also set their own learning goals, particularly achievement goals. Brookhart (2013b) discusses the relationships between students' achievement goals, motivation, and performance. She makes the case that classroom assessment evidence can be both the *goal* of motivated learning (e.g., a student wants to be able to turn in a high-quality lab report in a science class) and the *means* to that goal (therefore, the student works at learning the science content and the laboratory procedure; he or she sets a goal that is monitored during ongoing work toward the assessment). Part of the energy behind formative assessment's effects is derived from the simultaneous influence of classroom assessment on motivation and on achievement. The same evidence of where a student is going, where he or she is now, and what he or she should do next facilitates the student's cognition and at the same time supports motivation (self-agency) based on the feeling that what to do next is in sight and attainable. More research is needed on the relationship between unit-, lesson-, and task-specific goal setting by students and achievement, particularly since goals students set commit them to pursuing one particular outcome over another (Hadwin et al. 2011).

## 17.5 Phase Two: Progress Monitoring

A central purpose of both classroom assessment and self-regulation is to monitor learners' progress toward goals and provide feedback that can be used to deepen learning and improve performance. Monitoring progress toward goals can be a process of thinking about one's own thinking, or a related but distinct process of formatively or summatively evaluating the product-based evidence of learning against the standards for it. The former version of progress monitoring is known as metacognition and is largely internal to the learner. The latter version of progress monitoring is feedback and involves the solicitation of critiques from oneself and from others, often via classroom assessments.

Classroom assessment can support progress monitoring by addressing the three key questions identified by Hattie and Timperley (2007): Where am I going?, How

am I going?, and Where to next? Hattie and Timperley's model includes four types of feedback:

1. Task level: How well tasks are understood and performed.
2. Process level: The main processes needed to understand and perform tasks.
3. Self-regulation level: Self-monitoring, regulating, and directing of actions.
4. Self level: Personal evaluations of the learner.

They argue that self-level feedback (e.g., 'Good girl') is the least effective because it contains little task-related information. Feedback about processing and self-regulation are 'powerful in terms of deep processing and mastery of tasks,' and 'task feedback is powerful when the task information subsequently is useful for improving strategy processing or enhancing self-regulation (which it too rarely does)' (p. 91). Given what is known about how SRL is enhanced when learners receive feedback about strategy use (Zimmerman 2002), classroom assessments that provide process and self-regulation-level feedback could be quite effective in promoting both achievement and SRL. Feedback targeted at any level can come from a variety of sources, including students themselves, their peers, teachers, and technology.

### ***17.5.1 Self-Generated Feedback***

Good self-regulators evaluate their own performance and make adaptive attributions linked to deeper processing, better learning and achievement, positive affect, positive efficacy and expectancy judgments, persistence, and effort (Pintrich 2000). In a classroom context, this type of self-regulation can look like self-assessment, which is a process during which students reflect on the quality of their work, judge the degree to which it reflects explicitly stated goals or criteria, and revise their work accordingly (Andrade 2010). Self-assessment is a core element of self-regulation (Brown and Harris 2013) because it involves awareness of the goals of a task and checking one's progress toward them.

Brown and Harris's (2013) survey of research on self-assessment led them to conclude that there is evidence of a link between self-assessment and better self-regulation skills, 'provided such self-evaluation involves deep engagement with the processes affiliated with self-regulation (i.e., goal setting, self-monitoring, and evaluation against valid, objective standards)' (p. 386). For example, Panadero and his colleagues have explored the relationship between both task-level and process-level self-assessment and SRL in secondary students (Panadero et al. 2012) and undergraduates (Panadero et al. 2013, 2014; Panadero and Romero 2014). They used rubrics to scaffold self-generated task-level feedback and scripts (i.e., guides to the processes required by a task) for process-level feedback. The results suggest that, in general, students who engaged in self-assessment of their learning were more self-regulated, as measured by self-report questionnaires and/or think aloud

protocols, than were students in the comparison groups. Effect sizes were very small but statistically significant. Process-level self-assessment tended to be more closely associated with SRL than task-level self-assessment. This is probably the case because, as Brown and Harris put it, process-level self-assessment engages students with the processes affiliated with self-regulation. You get what you assess, as the saying goes.

Although many students probably do not explicitly self-assess in terms of task criteria, thereby missing an opportunity for self-regulated learning, the process is eminently teachable. Self-assessment has been investigated for its contribution to learning and performance in many contexts, including elementary and middle school writing, middle school mathematics, and high school social studies and technology. Andrade et al. (2008) had third and fourth graders read a model written assignment and generate a list of criteria as a class. Using rubrics based on those criteria, they self-assessed drafts of their stories and essays. Controlling for previous writing ability, the group that used the rubrics for self-assessment wrote better overall than a comparison group that self-assessed without formal criteria (Cohen's  $d = 0.87$ ). Andrade et al. (2010) replicated these findings with middle school students in fifth, sixth, and seventh grade (Cohen's  $d = 0.66$ ).

Ross et al. (2002) taught fifth- and sixth-grade students self-evaluation skills in mathematics, also using a method based on criteria. Their self-assessment instruction involved students in defining criteria, taught them how to apply the criteria, gave students feedback on these self-assessments against criteria, and helped students develop action plans based on the self-assessments. Controlling for previous problem-solving ability, students who self-assessed using criteria outscored a comparison group at solving mathematics problems.

Ross and Starling (2008) used the same four-component self-assessment training based on criteria with secondary students in a ninth-grade geography class. Students were learning to solve geography problems using global information systems (GIS) software, so the learning goals were about both accurate use of the software and using it to solve geography problems. Controlling for pretest computer self-efficacy, the treatment group outscored a comparison group on three different measures: production of a map using the software, a report explaining their problem-solving strategies, and an exam measuring knowledge of the mapping program. The largest difference was for the problem-solving explanations.

There is also limited evidence of a link between criteria-referenced self-assessment and self-efficacy, at least for girls. Self-efficacy, or the belief that one can succeed at a particular task (Bandura 2003), is a component of self-regulated learning that has an association with other motivational components of SRL such as task interest and persistence, as well as with relevant strategy use (Schunk and Usher 2011). Andrade et al. (2009) investigated the relationship between self-assessment according to a rubric and elementary and middle school students' ( $N = 268$ ) self-efficacy for writing. Students in the treatment group reviewed a model essay and



used a rubric to self-assess their drafts. Self-efficacy ratings were collected three times: before, during, and after writing the first draft. The results revealed interactions between gender and self-assessment: Average self-efficacy ratings increased as students wrote, regardless of gender or condition, but the increase in the self-efficacy of girls in the treatment group was much larger than the increase for girls in the comparison group (multivariate  $F(2, 169) = 3.61, p = 0.03$ ). There were no such differences for the boys (multivariate  $F(2, 99) = 0.07, p = 0.94$ ), suggesting that rubric-referenced self-assessment was associated only with the self-efficacy of girls. However, other studies found no clear relationship between self-assessment and self-efficacy (Meusen-Beekman et al. 2014; Panadero et al. 2013). This may be because other mediating variables exist, for example the degree to which students achieve their goals.

### ***17.5.2 Self- or Peer-Generated Feedback***

Students' peers can also play a role in progress monitoring. Meusen-Beekman et al. (2014) conducted a study of the relationship between self-regulated learning and peer or self-assessment with 695 sixth grade students in The Netherlands. Students in the treatment condition, which lasted 27 weeks, engaged in peer or self-assessment of three writing assignments. They also co-created the criteria for their writing tasks, set goals, made plans, and used checklists to monitor their progress. In these ways, the students provided themselves and each other with both task- and process-level feedback.

Analysis of the data from student self-report questionnaires, focus groups, and teacher observations suggest that the treatment had a statistically significant, positive association with self-regulation and intrinsic motivation, with no differences between the peer- and self-assessment conditions. Together with research done by Panadero and his colleagues on rubrics and scripts (Panadero et al. 2014), the results of Meusen-Beekman et al.'s (2014) study support claims that formative peer and self-assessment can scaffold self-regulation, particularly when the feedback received from either source is focused on both the criteria for the task at hand and the processes employed to produce work that meets them. It might be important to note, however, that the control condition in Meusen-Beekman et al.'s study did not allow for revision, which could have suppressed students' self-regulation and motivation.

Similarly, Graham et al. (2012) found that involving students in prewriting activities, peer assistance, clarifying goals, and assessment with feedback was an important series of writing interventions that raised writing achievement. Graham et al. called this kind of intervention 'scaffolding writing' (p. 887), but it also may be described as formative assessment.

### ***17.5.3 Interpretation of Feedback During Progress Monitoring***

We have long known that the action taken by a learner in response to feedback depends, in part, on the way in which it was received (Black and Wiliam 1998), because a learner's response to feedback involves interpretation of that feedback. Most research on the nature of learners' interpretations of feedback has focused on the effects of feedback on affect, particularly motivation (Brookhart 2013b). Empirical studies of the effects of students' interpretations of feedback on learning and achievement are scarce.

Draper (2009) developed a theoretical argument that stresses how students' interpretations of ambiguous feedback determine whether that feedback is useful or not. He postulates at least six possible interpretations of feedback:

1. Technical knowledge or method (e.g., concluding that one did not use the best information or method to complete the task, both of which can be improved).
2. Effort (e.g., deciding that one did not leave enough time to do a task well).
3. Method of learning about a task (e.g., realizing that one did not seek out the right information about the task, or did not understand the criteria for the task).
4. Ability (e.g., believing that one does not have the necessary aptitude to succeed at a task).
5. Random (e.g., assuming nothing was done incorrectly so success is possible next time without adjustment or revision).
6. The judgment process was wrong (e.g., determining that the feedback was incorrect).

It is very likely that students' self-regulatory responses to feedback are determined by the type of interpretation they make of a given instance of feedback. Research that tests this or related theories, and the ways in which classroom assessment can influence students' interpretations of feedback and subsequent attempts to regulate their learning, is needed.

### ***17.5.4 Feedback from Grades***

Some students are very effective self-regulators, and there is evidence that these students use feedback from all sources, including grades, for specific information about the content of an assessment as well as for general information about how to study or do project work better. Brookhart (2001) interviewed successful students in high school English and Anatomy classes to learn their perspectives on the formative and summative aspects of classroom assessments. These students, mostly

from honors classes in a well-resourced high school, were taking challenging classes and were invested in their education and in getting good grades. Students were asked about specific graded tests or assignments. A striking finding was that they considered all assessment to be formative to some degree. They considered studying for a test or doing a project as a contribution to their learning. They looked for ways to transfer their current learning to future study. They intentionally worked at self-monitoring, reflecting on ‘how well they did,’ a phrase they used to mean both the grade they received and what they thought they learned.

Although high-achieving students like those in Brookhart’s (2001) study report using formal evaluations for the purposes of progress monitoring, too few students actually do so. In fact, some very unsuccessful students also use summative outcomes to regulate their learning in unhelpful ways by developing learned helplessness (Dweck 1976). Fortunately, scaffolding can be put in place to help students use grades or scores to monitor their progress. For instance, Brookhart et al. (2004) studied third graders learning their 0–9 multiplication facts. Every week for ten weeks, they took a 100-fact, 5-minute timed test. Each week they predicted what their next score would be, and then graphed their actual score next to it, using a bar graph. At the time of prediction they also used a reflection sheet to set a learning goal for the next week (e.g., ‘do the 8 tables better’) and plan a strategy for reaching that goal (‘practice with flash cards’). The reflection sheet led them to set the goal and strategy based on how they thought their previous week’s goal and strategy had worked.

An analysis of the students’ reflections showed that most students expressed a mastery goal orientation. Students learned their multiplication facts quickly and enjoyed the reflection, especially graphing their ‘steps’ (their grades) each week. Students who achieved 100 % before the ten weeks challenged themselves to do the test in four and then three minutes rather than stop the project. This project combined features of formative and summative assessment; students tracked their progress and used the results formatively, but their grades were also derived from their performance on the timed tests.

If students are to use grades to monitor their achievement, then those grades must reflect meaningful standards of learning and students’ progress toward them (Guskey 2009). For the past 10 years or so, a movement known as standards-based grading has been gaining momentum in U.S. schools. Teachers using traditional grading practices often combine appraisals of effort and behavior, as well as learning, into a grade. In contrast, teachers who employ standards-based or learning-focused grading assess student work in terms of achievement alone and report measures of effort and behavior separately. Grades should be useful for the progress-monitoring phase of self-regulation of learning. If students are to use their grades in an evidentiary process to regulate ever more learning, the grades need to be evidence of having learned or not learned.

## 17.6 Phase Three: Revision and Adjustment

In terms of self-regulated learning, revision and adjustment refer to the types of cognitive and metacognitive activities in which students engage in order to adapt and change their thinking, including selecting and using cognitive strategies (Pintrich 2000). Attribution of success and failure is another aspect of this phase (Pintrich and Zusho 2002). Attributions are both cognitive and motivational in nature, and linked to academic achievement. Feedback and evaluation from any source can affect students' attributions (Oren 2001). For instance, Dweck (2006) has shown that teachers can change the way children come to understand their abilities related to an activity simply through the choice of feedback they offer in moment-to-moment feedback: Praising students for their intelligence (e.g., 'You are so smart') tends to induce a fixed mindset, while praise focused on effort or process (engagement, perseverance, effective strategy use, or improvement, e.g., 'You worked hard to improve this') fosters a growth mindset.

From a classroom assessment perspective, the revision and adjustment phase of learning can involve revision of student work, particularly after receiving feedback. We know very little about the adjustments to goal-directed action that students make in light of classroom assessment. This lack of information about what students actually do in response to feedback reflects the fact that research has tended to employ measures of outcomes and products rather than of the processes of learning and revision. Research is needed on the adjustments that students make to their work and learning processes (if any) in response to both formative and summative assessment.

One issue on which there is consensus is that if feedback is to be useful, it must be focused on criteria, describe reasonable next steps, and followed by opportunities to close the gap between current and desired performance through retakes or revision (Andrade 2010; Boud 2000; Brookhart 2013a; Nicol and Macfarlane-Dick 2006). Unfortunately, teachers often move on to the next topic or assignment, citing time pressures that prevent resubmission after feedback. When self- or peer-generated feedback followed by revision is part of a regular class routine, however, students share the feedback burden with teachers (Lipnevich et al. 2014), and the likelihood of self-regulation is greater.

## 17.7 Implications for Classroom Practice

Ample research has shown that supporting students in learning to use self-regulation strategies is related to subsequent improvements in academic achievement, especially when instruction in SRL begins in the late childhood or early adolescent years (de Boer et al. 2012). Given the similarities between classroom assessment and SRL, and burgeoning evidence of an influence of the former on the latter, an obvious practical implication is to use assessments, especially

formative assessments, as a form of SRL instruction. This might be a simple matter of framing: Rather than telling students to peer or self-assess in order to get a better grade, we can explain that seeking feedback from oneself and others is a learning skill that, when honed into a habit, is a hallmark of successful learners.

Clear learning goals and criteria are the foundation on which both formative assessment and SRL rest. Students cannot accurately evaluate their progress without an understanding of the standards held by their teacher (Allal 2010; McMillan 2011; Nicol and Macfarlane-Dick 2006; Stiggins 2008). Educators are beginning to grasp the importance of communicating learning goals to their students, but anecdotal evidence from professional development work, as well as some research (Antoniou and James 2014; Saito and Inoi 2015; Wylie and Lyon 2015), suggests that teachers have a much harder time conceptualizing and communicating task-specific success criteria. From the point of view of a student who has not yet reached a learning goal, however, a goal without criteria is not very useful. A clear implication for practice is for teachers to develop skills in conceptualizing, communicating, and using success criteria.

Learning goals and success criteria are not enough, however. Another clear implication for practice is to employ assessments that present students with process- and self-regulation-oriented feedback. Findings by Panadero and colleagues (Panadero et al. 2012, 2013, 2014) suggest that process-oriented scripts tend to be more highly associated with SRL than rubrics. These findings support Hattie and Timperley's (2007) claims regarding the power of feedback that informs students about how to effectively engage in tasks and how to monitor and regulate their progress. A related implication is that opportunities to revise are essential if assessment is to lead to self-regulated learning.

A less obvious practical implication of our discussion of the relationship between assessment and SRL is the need to carefully scaffold constructive interpretations of feedback and attributions of success or failure. We cannot assume that students always eagerly receive information about their achievement, whether formative or summative, and happily apply it in ways that deepen their learning and improve their products. Tools and procedures are needed that increase the likelihood of interpretations of feedback that result in beneficial self-regulatory responses.

## 17.8 Challenges of Implementation

Classroom assessment plays a pivotal role in student goal setting, progress monitoring, and revision and adjustment. As the literature shows, assessment can support the self-regulation of learning in classroom settings if it provides students with ways to participate in all three phases with intentionality and ownership. The literature has also identified a major challenge for implementation, namely, a typical classroom environment focused on grading rather than learning that sets up assessments as trials that only some can win (Covington 1992). A related

challenge is the need for teachers to turn conventional instructional and assessment planning on its head and approach these tasks from the students' point of view (Andrade 2010). By so doing, teachers can maximize the likelihood that students will have the opportunities and tools they need to take ownership of their own learning by self-regulating it.

To meet these challenges, teachers and administrators will need assessment literacy, including understanding the purposes of assessment, the value of it, and effective, student-centered classroom practices. Teachers often use assessment tasks designed by someone else without a deep understanding of the content or the reasons for the design choices. This could limit the kind of process or SRL feedback they are able to give to students. In addition, even assessment for learning can end up being teacher centered rather than student centered (Jonsson et al. 2015), or, through teacher misconceptions, can end up not being formative at all. For example, in one study, 20 % of reported 'formative assessment' was not, in fact, formative (e.g., giving points for a participation grade to students who answered random questions; Wylie and Lyon 2015). When such things happen, students experience an evaluative rather than a learning-focused classroom environment. Assessment literacy, including a deep understanding of the formative assessment process and students' and teachers' roles in it, is needed in order to overcome these pitfalls.

Promoting assessment literacy requires a two-pronged approach in which teachers learn, and have the opportunity to apply, sound assessment strategies, not only in workshops but in their regular classrooms, and at the same time work to empower students as owners of their own learning who are capable of, and practice, self-assessment. This kind of professional development takes time, requires participatory professional development techniques, and requires modeling the same kind of assessment teachers need to practice.

Developing, communicating, and using success criteria with students is one of the central aspects of formative assessment, but teachers are much better at describing learning goals for students than articulating the criteria that indicate deep learning or high quality work (Antoniou and James 2014; Wylie and Lyon 2015). Professional development that focuses on success criteria in the context of formative assessment is imperative. The authors' experiences suggests one good way to help teachers develop success criteria that are about learning, rather than about the requirements for the task, involves analyzing good examples of student work, and critiquing and revising poor examples.

Another challenge is related to the need for assessments that focus on process. The development of self-regulated learning benefits from feedback about strategy use and the benefits of using them (Zimmerman 2002), but teachers tend to provide feedback about performance, not processes. Students need feedback about both performance and process (Hattie and Timperley 2007; Panadero et al. 2012). Teachers have a difficult time giving descriptive feedback in time for students to extend their learning and amend their performances, instead of giving feedback that explains summative evaluations (Wylie and Lyon 2015). The provision of effective, process-oriented feedback is an aspect of assessment for learning that needs to be addressed head-on in professional development. Professional development that

proceeds from analyzing examples of others' feedback through practicing and analyzing one's own feedback can be effective.

Finally, there is the inexorable pull of standardized tests in the U.S. context (Berliner and Nichols 2007) and elsewhere (e.g., Scotland; Hayward 2015), including in Confucian-heritage settings with traditions of high-stakes examinations (Carless 2011; Ratnam-Lim and Tan 2015). Teachers' and students' beliefs about the importance of learning and students' self-regulation of learning must be very robust indeed to stand up to the pressure of high-stakes examinations. Classroom assessment that facilitates learner autonomy and self-regulation has never been more needed than now, given the current prominence of standardized tests.

## 17.9 Conclusion

Covington (1992) reviewed several decades' worth of literature to demonstrate that all students, especially lower achievers, can be successful learners if they are given opportunities to understand their own learning, set their own goals, pursue appropriately leveled tasks, and receive feedback targeted to their needs. The key is students having some control over what and how they learn and receiving sufficient information to strengthen their effort-outcome beliefs and thus their attribution of success to their own efforts. Classroom assessment can play a large part in creating the conditions for student agency identified by Covington. In fact, because assessment is an integral part of most classroom activities, and SRL strategy instruction is most effective when it is embedded in an authentic learning context (Paris and Paris 2001), classroom assessment presents a unique opportunity to support students in becoming successful learners—if it is intentionally used to do so.

## References

- Allal, L. (2010). Assessment and the regulation of learning. In P. Peterson, E. Baker, & B. McGraw (Eds.), *International encyclopedia of education* (Vol. 3, pp. 348–352). Oxford: Elsevier.
- Andrade, H. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 57(5), 13–18.
- Andrade, H. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. Andrade & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 90–105). New York: Routledge.
- Andrade, H., Du, Y., & Mycek, K. (2010). Rubric-referenced self-assessment and middle school students' writing. *Assessment in Education: Principles, Policy and Practice*, 17(2), 199–214.
- Andrade, H., Du, Y., & Wang, X. (2008). Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practices*, 27(2), 3–13.
- Andrade, H., Wang, X., Du, Y., & Akawi, R. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *The Journal of Educational Research*, 102(4), 287–302.

- Antoniou, P., & James, M. (2014). Exploring formative assessment in primary school classrooms: Developing a framework of actions and strategies. *Educational Assessment, Evaluation and Accountability*, 26(2), 153–176. doi:10.1007/s11092-013-9188-4.
- Bandura, A. (2003). *Self-efficacy: The exercise of control*. New York, NY: W.H. Freeman and Company.
- Belgrad, S. (2013). Portfolios and e-portfolios: Student reflection, self-assessment, and goal setting in the learning process. In J. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 331–346). Thousand Oaks, CA: SAGE.
- Berliner, D., & Nichols, S. (2007). *Collateral damage: How high-stakes testing corrupts America's schools*. Cambridge, MA: Harvard Education Press.
- Black, P., McCormick, R., James, M., & Pedder, D. (2006). Learning how to learn and assessment for learning: A theoretical inquiry. *Research Papers in Education*, 21(2), 119–132. doi:10.1080/02671520600615612.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, 5, 7–74.
- Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research*, 31, 445–457.
- Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151–167.
- Brookhart, S. M. (2001). Successful students' formative and summative use of assessment information. *Assessment in Education: Principles, Policy and Practice*, 8, 153–169.
- Brookhart, S. M. (2013a). *How to create and use rubrics for formative assessment and grading*. Alexandria, VA: ASCD.
- Brookhart, S. M. (2013b). Classroom assessment in the context of motivation theory and research. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 35–54). Thousand Oaks, CA: SAGE.
- Brookhart, S., Andolina, M., Zuzza, M., & Furman, R. (2004). Minute math: An action research study of student self-assessment. *Educational Studies in Mathematics*, 57(2), 213–227. doi:10.1023/B:EDUC.0000049293.55249.d4.
- Brown, G. T., & Harris, L. R. (2013). Student self-assessment. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 367–393). Thousand Oaks, CA: SAGE.
- Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, 79, 474–482.
- Butler, D. L., & Cartier, S. (2004). Promoting students' active and productive interpretation of academic work: A key to successful teaching and learning. *Teachers College Record*, 106(9), 1729–1758.
- Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology*, 78, 210–216.
- Butler, D. L., & Schnellert, L. (2015). Success for students with learning disabilities: What does self-regulation have to do with it? In T. Cleary (Ed.), *Self-regulated learning interventions with at-risk youth: Enhancing adaptability, performance, and well-being being* (pp. 89–112). Washington, DC: APA Press.
- Carless, D. (2011). *From testing to productive student learning: Implementing formative assessment in Confucian-heritage settings*. Abingdon: Routledge.
- Covington, M. V. (1992). *Making the grade: A self-worth perspective on motivation and school reform*. Cambridge: Cambridge University Press.
- de Boer, H., Donker, A. S., & Kostons, D. D. N. M. (2012). *Effective strategies for self-regulated learning: A meta-analysis*. Groningen: GION (Gronings Instituut voor Onderzoek van Onderwijs), Rijksuniversiteit Groningen.
- Draper, S. (2009). What are learners actually regulating when given feedback? *British Journal of Educational Technology*, 40(2), 306–315.
- Dweck, C. (2006). *Mindset: The new psychology of success*. New York: Random House.



- Dweck, C. S. (1976). Children's interpretation of evaluative feedback: The effect of social cues on learned helplessness. *Merrill-Palmer Quarterly*, 22(2), 105–109.
- Graham, S., Kihara, S., McKeown, D., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology*, 104, 879–896.
- Guskey, T. R. (Ed.). (2009). *Practical solutions for serious problems in standards-based grading*. Thousand Oaks, CA: Corwin Press.
- Hadwin, A., Järvelä, S., & Miller, M. (2011). Self-regulated, co-regulated, and socially shared regulation of learning. In B. Zimmerman & D. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 65–86). New York: Routledge.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81–112.
- Hayward, L. (2015). Assessment is learning: The preposition vanishes. *Assessment in Education: Principles, Policy & Practice*, 22(1), 27–43. doi:10.1080/0969594X.2014.984656.
- Heritage, M. (2010). *Attributes of effective formative assessment*. Washington, DC: Council of Chief State Officers.
- Jonsson, A., Lundahl, C., & Holmgren, A. (2015). Evaluating a large-scale implementation of assessment for learning in Sweden. *Assessment in Education: Principles, Policy and Practice*, 22(1), 104–121.
- Lipnevich, A., & Smith, J. (2008). *Response to assessment feedback: The effects of grades, praise, and source of information*. Research report RR-08-30. Princeton, NJ: Educational Testing Service.
- Lipnevich, A. A., McCallen, L. N., Miles, K. P., & Smith, J. K. (2014). Mind the gap! Students' use of exemplars and detailed rubrics as formative assessment. *Instructional Science*, 42(4), 539–559.
- McMillan, J. (2011). *Classroom assessment: Principles and practice for effective standards-based instruction*. New York: Pearson.
- Meusen-Beekman, K., Joosten-ten Brinke, D., & Boshuizen, H. (2014). *The effects of formative assessment on self-regulated learning skills by sixth grade pupils*. Paper presented at the Conference of EARLI SIG1: Assessment and Evaluation, Madrid, Spain.
- Moss, C., Brookhart, S., & Long, B. (2013). Administrators' roles in helping teachers use formative assessment information. *Applied Measurement in Education*, 26, 205–218.
- Nicol, D., & 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.
- Oren, D. (2001). Evaluation systems and attributional tendencies in the classroom: A sociological approach. *Journal of Educational Research*, 76(5), 307–312.
- Panadero, E., Alonso-Tapia, J., & Huertas, J. A. (2012). Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and Individual Differences*, 22(6), 806–813. doi:10.1016/j.lin-dif.2012.04.007.
- Panadero, E., Alonso-Tapia, J., & Huertas, J. A. (2014). Rubrics vs. self-assessment scripts: Effects on first year university students' self-regulation and performance. *Journal for the Study of Education and Development*, 3(7), 149–183. doi:10.1080/02103702.2014.881655
- Panadero, E., Alonso-Tapia, J., & Reche, E. (2013). Rubrics vs. self-assessment scripts effect on self-regulation, performance and self-efficacy in pre-service teachers. *Studies in Educational Evaluation*, 39(3), 125–132. doi:10.1016/j.stueduc.2013.04.001.
- Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice*, 21(2), 133–148. doi:10.1080/0969594X.2013.877872.
- Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, 36(2), 89–101.

- Pintrich, P., & Zusho, A. (2002). The development of academic self-regulation: The role of cognitive and motivational factors. In J. Eccles & A. Wigfield (Eds.), *Development of achievement motivation* (pp. 249–284). San Diego, CA: Academic Press.
- Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, 92, 544–555.
- Ratnam-Lim, C. T. L., & Tan, K. H. K. (2015). Large-scale implementation of formative assessment practices in an examination-oriented culture. *Assessment in Education: Principles, Policy & Practice*, 22(1), 61–78. doi:10.1080/0969594X.2014.1001319
- Ross, J. A., Hogaboam-Gray, A., & Rolheiser, C. (2002). Student self-evaluation in grade 5-6 mathematics: Effects on problem-solving achievement. *Educational Assessment*, 8, 43–58.
- Ross, J. A., & Starling, M. (2008). Self-assessment in a technology-supported environment: The case of grade 9 geography. *Assessment in Education: Principles, Policy and Practice*, 15(2), 183–199.
- Saito, H., & Inoi, S. (2015). *Junior and senior high school EFL teachers' use of formative assessment: A mixed method study*. Paper presented at the Language Testing Research Colloquium, Toronto.
- Schunk, D., & Usher, E. (2011). Assessing self-efficacy for self-regulated learning. In B. Zimmerman & D. Schunk, (Eds.), *Handbook of self-regulation of learning and performance* (pp. 282–297). New York: Routledge.
- Stiggins, R. (2008). *An introduction to student-involved assessment FOR learning*. Upper Saddle River, NJ: Pearson Education.
- Torrance, H., & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal*, 27(5), 615–631.
- Winne, P. (2011). A cognitive and metacognitive analysis of self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performances* (pp. 15–32). New York: Routledge.
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ: Lawrence Erlbaum.
- Wylie, E. C., & Lyon, C. J. (2015). The fidelity of formative assessment implementation: Issues of breadth and quality. *Assessment in Education: Principles, Policy & Practice*, 22(1), 140–160. doi:10.1080/0969594X.2014.990416.
- Zimmerman, B. (2011). Motivational sources and outcomes of self-regulated learning and performance. In B. Zimmerman & D. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 49–64). New York: Routledge.
- Zimmerman, B., & Schunk, D. (Eds.). (2011). *Handbook of self-regulation of learning and performance*. New York: Routledge.
- Zimmerman, B. J. (2002). Achieving self-regulation: The trial and triumph of adolescence. In F. Pajares & T. Urdan (Eds.), *Academic motivation of adolescents*. Greenwich, CT: Information Age Publishing.

# Chapter 18

## Scaffolding Self-Regulated Learning Through Self-Assessment and Peer Assessment: Guidelines for Classroom Implementation

Ernesto Panadero, Anders Jonsson and Jan-Willem Strijbos

**Abstract** Although the focus on feedback and student involvement in Assessment for Learning (AfL) appears to align very well with theories of Self-Regulated Learning (SRL), and also seems to be the main reason for many researchers' interest in formative assessment, the actual relationship between AfL and SRL is an issue of debate. In this chapter, we therefore explore the relationship between two AfL practices, namely, self-assessment and peer assessment, and SRL. These AfL practices emphasize student feedback and are both thought to increase student involvement in assessment. They also have evident connections to SRL models of self-regulation and co-regulation. Special attention is given to strategies for the implementation of peer and self-assessment in the classroom. In particular, guidelines are presented on teachers' mediating and modeling role in peer and self-assessment, as well as on how to use formative assessment instruments, such as rubrics, scripts, and prompts, in order to promote student involvement in assessment.

### 18.1 Introduction

The promotion of students' active involvement in assessment is an integral part of Assessment for Learning (AfL). Still, in studies on the implementation of AfL it has been noted that teachers, when given the choice, may choose *not* to involve students in the assessment process. Instead they may focus primarily on other AfL practices, such as clarifying and sharing assessment criteria, designing learning situations, and

---

E. Panadero (✉)  
Universidad Autonoma de Madrid, Madrid, Spain  
e-mail: ernesto.panadero@uam.es

A. Jonsson  
Kristianstad University, Kristianstad, Sweden

J.-W. Strijbos  
Ludwig-Maximilians-Universität München, Munich, Germany

providing external feedback (Jonsson et al. 2015). This is problematic, not only because the AfL practices then run the risk of becoming more teacher- and teaching-centered rather than student- and learning-centered; students may thus lose an important opportunity to develop the capacity to self-regulate their learning.

The relationship between AfL and Self-Regulated Learning (SRL) has been a topic of interest and debate among AfL researchers since the beginning of the field. For instance, the seminal review by Black and Wiliam (1998) included articles exploring this relationship (e.g., the review by Butler and Winne (1995) that connects SRL and feedback). More recently, Nicol and McFarlane-Dick (2006) proposed seven principles of ‘good feedback practice’ that are thought to promote SRL. Likewise, Clark (2012) discussed, from a theoretical point of view, how formative feedback may actualize and reinforce SRL strategies among students. These contributions, however, give insufficient attention to what we know about SRL theories and models (Panadero and Alonso-Tapia 2013; Zimmerman and Schunk 2001). Most importantly, they do not explain in any detail how AfL may affect the different components of the SRL processes. This means that the relationship between AfL and SRL is more often assumed than explicitly supported by findings from empirical research.

In this chapter, we will therefore present research that has empirically investigated the relationship between AfL practices and SRL. Specifically, we will start from a key SRL model and explore how its components may be affected by two AfL practices that, in our view, are highly important in the context of SRL, namely self-assessment (SA) and peer assessment (PA). By presenting examples from research on SA and PA, we will illustrate what is currently known about these practices in relation to SRL. From the same research base, we will also present recommendations on how to facilitate the implementation of SA and PA in the classroom.

## 18.2 A Brief Overview of SRL

According to Zimmerman (2000), SRL can be defined as ‘self-generated thoughts, feelings, and actions that are planned and cyclically adapted to the attainment of personal goals’ (p. 14). SRL has become one of the main theories in educational psychology and it is often cited as the core competence required for ‘learning to learn’ (Zimmerman and Schunk 2011). There are a number of SRL models from different theoretical perspectives (Panadero and Alonso-Tapia 2014), but they all highlight the following common characteristics—namely that SRL is (a) a cyclical process, (b) comprised of cognition, metacognition, motivation and emotion, and (c) a set of skills that can be developed and learned.

One of the primary reasons for the growing interest in SRL research is the impact that SRL has on student learning and performance. A number of studies have shown that SRL skills are important for the success of students (Dignath and Büttner 2008).

Also, acquiring SRL skills during one's education may impact subsequent 'lifelong learning' (e.g., de la Harpe and Radloff 2000).

There is an increasing amount of evidence that SRL skills can be learned and even taught as a set of generic skills. A number of researchers have conducted interventions in classrooms in order to develop students' SRL skills, either working with the students (Dignath and Büttner 2008) and/or with the teachers (Kramarski and Michalsky 2009). Working with teachers by including them in the interventions is crucial since they have an important role in shaping the classroom learning environment (Boekaerts and Corno 2005). One of the critical aspects in this respect is classroom assessment. Depending on the existing assessment culture, students are likely to make use of different approaches to learning, focusing mainly on strategies for either deep or surface learning (e.g., Segers et al. 2006), which highlights the interplay of classroom assessment practices (AfL) and student learning strategies (e.g., deep vs. surface learning, SRL).

### 18.3 The Relationship Between AfL and SRL

Some of the earliest work on the relationship between AfL and SRL dates back to the 1990s. Butler and Winne (1995) analyzed the relationship between feedback and SRL, establishing the distinction between internal and external feedback. This is important because it emphasizes that teachers' external feedback impacts students' development of SRL skills via their internal feedback. In their review of classroom assessment, Black and Wiliam (1998) explicitly connected SRL to classroom assessment and formative assessment practices, thus laying the foundation for what was later to become known as AfL.

Whereas previous research on SRL was more theoretically oriented as compared to research on AfL, which tended to be more practical and classroom focused, there has been a shift of late in the literature towards a more balanced use of both theoretical contributions and empirical data in both fields (e.g., Andrade and Brookhart 2014; Panadero 2011; Winne 2014). As pointed out by Wiliam (2014), a benefit of the SRL perspective for AfL is that it allows practical classroom techniques to be theorized and more easily shared; at the same time, AfL practices may enhance students' SRL skills by providing students with the opportunity to practice these skills and by providing (external) feedback that can support student learning. Two AfL practices that can affect SRL—through their emphasis on student involvement and feedback—are SA and PA. In the next two sections, we therefore analyze the relationship between SA and SRL theory—more specifically, the model proposed by Zimmerman and Moylan (2009)—and the relationship between PA and co-regulation or socially shared regulation of learning (Hadwin et al. 2011; McCaslin 2009). In addition to establishing theoretical connections, we analyze empirical evidence that supports such connections.

## 18.4 Self-Assessment

### 18.4.1 Definition

According to the frequently cited definition of Boud and Falchikov (1989), SA occurs when students make ‘judgements about their own learning, particularly about their achievements and the outcomes of their learning’ (p. 529). SA may involve a wide range of activities, from asking students to grade their own work without further reflection (i.e., self-grading/self-marking), at one end of the spectrum, to having them make comprehensive analyses of their own performance on complex tasks (Panadero et al. 2016), at the other end. SA has been shown to have positive effects on student performance with a median effect size (Cohen’s *d*) between 0.40 and 0.45 (Brown and Harris 2013), which is consistent with the effects reported by Boud and Falchikov (1989).

### 18.4.2 SA as an Instructional Approach and an SRL Component

There seems to be two lines of SA research: one coming from a more teacher-centered perspective linked to AfL and the other coming from the SRL literature (Olinia and Sullivan 2004; Panadero and Alonso-Tapia 2013). Whereas SRL research is based on an ‘internal perspective’ of SA and focuses primarily on the ‘inner processes’ and on self-regulation as a generic skill, AfL research considers SA often as a context-dependent skill that is not easily transferable across different situations or subjects. Although these differences have become less pronounced in recent years (e.g., Kostons et al. 2012), some researchers (e.g. Panadero and Alonso-Tapia 2013) still argue that it is important to differentiate between AfL and SRL approaches to SA and how they emphasize different aspects of SA because this will support a more coherent use of AfL practices based on at their enhancement of SRL.

Interestingly, the potential of merging the AfL and the SRL approaches with respect to SA did not attract much attention until recently when there has been an increased interest in this topic (Andrade and Brookhart 2014; Kostons et al. 2012; Panadero and Alonso-Tapia 2013). As proposed by Wiliam (2014), research in this area needs to build on the strengths of each of these approaches in order to bridge theory and practice. The main idea is that when a teacher provides the space for working with SA in his/her classroom, this can improve students’ capacity to self-assess their own work and thereby improve their SRL skills.

### **18.4.3 Zimmerman's SRL Model and SA**

Zimmerman's cyclical model of self-regulation is one of the most cited models in the SRL literature (Panadero and Alonso-Tapia 2014; Zimmerman 2013). Zimmerman's model (see Zimmerman and Moylan 2009) consists of three cyclical phases: (a) *Forethought* which includes task analysis and self-motivation beliefs, (b) *Performance* which includes self-control and self-observation processes, and (c) *Self-reflection* which includes self-judgment and self-reaction. It is important to note that the different SRL phases are not closed and have a recursive nature, meaning that the self-reflection phase results will have effects in the forethought phase the next time the student performs the task. SA can take place during all phases of the model (Panadero and Alonso-Tapia 2013). In the next sections we discuss the phases of the Zimmerman model in relation to AfL research.

#### **18.4.3.1 Forethought Phase**

Researchers often recommend preparing for SA as early as possible when planning an activity (Andrade and Valtcheva 2009; Panadero and Alonso-Tapia 2013; Topping 2003). As visualized in the forethought phase of the model, students analyze the task, set goals, and identify which strategies are needed. According to research in AfL, it may be beneficial for teachers to discuss the assessment criteria with the students before starting the activity. Students can then use these criteria to set more realistic goals for the activity and to evaluate their work both during the process and afterwards (Panadero and Alonso-Tapia 2013). If the students do not have access to the criteria until after they have performed the task, they are likely to discover some aspects of their performance that they should have included from the beginning, but by then it might already be too late. A particular branch of AfL research emphasizing the need for clear criteria is research on scoring rubrics used by students (Jonsson and Svingby 2007; Panadero and Jonsson 2013). In this line of research, it is a common recommendation to provide the students with explicit assessment criteria before performing the task and in some cases even to negotiate the criteria with the students. With the aid of explicit criteria, students are thought to become more motivated to perform the task and also able to set more realistic goals for themselves (Andrade and Du 2005; Jonsson 2014; Panadero and Alonso-Tapia 2013).

#### **18.4.3.2 Performance Phase**

During the performance phase, students have to assess how well they are progressing towards the goals they established in the planning phase. This activity is known as 'metacognitive monitoring' in the model, as it involves reflection on one's own work. Emotional aspects of performance are also considered (for example, interest incentives which are related to motivation and emotion). During this phase there are

different ways in which AfL practices may influence the SRL cycle. First, if assessment criteria are known to the students at the beginning of the task, it will be easier for them to check whether they are ‘on track.’ Second, if students know that they will receive formative feedback (i.e., feedback that is aimed at enhancing their learning), they are less likely to feel anxious or stressed during the performance, as they know there will be opportunities to improve. And, third, students who have become used to receiving detailed feedback will likely be more motivated to ask for help if they get stuck and to activate SRL strategies in order to overcome the challenges encountered.

### **18.4.3.3 Self-Reflection Phase**

During the self-reflection phase, the students’ main focus is on evaluating their own work (which is facilitated via external and/or internal feedback; Butler and Winne 1995) and they make inferences about the causes for success or failure. There are several ways in which AfL practices might impact this phase. First, with access to explicit assessment criteria, students may make a more valid assessment of their work as they know the key aspects expected in the final product. Second, they may thus make more accurate interpretations of reasons for success or failure. When students understand the reasons for a weak performance, they can more easily attribute their level of performance to factors that they can potentially influence (Panadero and Alonso-Tapia 2013; Zimmerman and Moylan 2009). In this situation, students’ reactions are less likely to affect self-image, motivation, and learning strategies in a way that is detrimental for their learning. Third, since the feedback in AfL practices is oriented towards promoting learning, students can use the feedback in order to improve their performance, especially if they are given the opportunity to revise their work. This is one of the key findings in the research on students’ use of feedback: while it is well established that many students do not use the feedback they receive, most students do so if the use of feedback is an integrated part of instruction, which allows students to revise their work or to perform a similar task assessed with the same criteria (Jonsson 2013).

## ***18.4.4 Empirical Evidence of the Relationship Between SA and SRL***

A significant number of papers advocate a theoretical relationship between SA and SRL (Lan 1998; Paris and Paris 2001) and even some empirical papers try to establish such a relationship based on teachers’ and students’ perceptions of assessment practices (Harris and Brown 2013; Tan 2012). However, studies with experimental or quasi-experimental designs, where the effects of SA treatments on SRL are investigated and compared against control groups, are quite scarce.



Nevertheless, several studies do exist and they show the benefits of implementing SA interventions (as proposed by AfL) for the enhancement of SRL skills. Two lines of research in this area can be distinguished.

In the first line of research, Kostons et al. (2012) explored two aspects of SA with secondary school students (age  $M = 15.23$ ). First, they investigated whether SA skills and ‘task-selection accuracy’ (i.e., the extent to which students are able to choose a task according to their own capability) could be acquired either by observing a model performing the task or by repeated task practice. Second, they investigated if SA training would have an impact on students’ SRL skills. One of their main findings was that students could be trained in SA via observing a model performing the task. Furthermore, students who were trained to self-assess in combination with training in task-selection accuracy outperformed the control group on SRL effectiveness, as measured by student responses on mental effort and SA rating scales. However, only students who observed a model (and not students who were exposed to repeated training) outperformed the control group on SA accuracy. The results from this study have important implications for the field. First and foremost, SA training may indeed enhance the use of SRL skills. Another important implication, although this may seem obvious, is that not all SA interventions have the same effect. Furthermore, the results indicate that the accuracy of student SA may not be vital for improving students’ SRL skills through SA training.

In the second line of research, Panadero and colleagues explored in a series of studies on how SA, as promoted via scripts and rubrics, influenced the use of SRL skills (Panadero 2011). A script is a list of specific questions, structured in steps that model how an expert in the field would approach a complex task from beginning to end. Scripts can be used as way to scaffold students’ strategies and thinking when solving complex tasks. Rubrics are an assessment instrument that specifies which aspects of student performance are to be assessed and provides descriptions of different levels of quality for each aspect. An essential difference between scripts and rubrics is that scripts are designed as instructional resources, whereas rubrics are primarily assessment instruments. Rubrics therefore do not include instructions about how to solve a task but stress how to *evaluate* either the process or the product (or both).

In the first of their studies with secondary school students (age  $M = 15.9$ ), Panadero et al. (2012) found that the level of SRL in the script group, as measured through think-aloud protocols, was higher than in the rubric and control groups. Additionally, the rubric group showed a higher level of SRL than the control group. One important conclusion is therefore that SA training via either a rubric or a script may enhance the use of SRL strategies. The self-reported SRL data did not show significant differences<sup>1</sup>: students reported similar levels of SRL, while in practice there were significant differences in their use of SRL strategies.

---

<sup>1</sup>There were significant differences on the performance/avoidance SRL scale but it was in interaction with the type of feedback and type of instructions. Therefore the effect occurred in interaction with other variables that will not be discussed here.

The comparison script versus rubric was further analyzed in two subsequent studies using only self-reported data. In Panadero et al. (2013) it was found that preservice teachers using scripts scored higher on ‘Learning SRL’ (positive type of SRL), whereas students using rubrics decreased their scores on ‘Performance/avoidance SRL’ (negative type of SRL). In a similar fashion, Panadero et al. (2014a) found that first year psychology students who used rubrics decreased Learning SRL, whereas their use of scripts increased Performance/avoidance SRL. Therefore, scripts seem to have a positive effect on Learning SRL, enhancing it more than do rubrics, while rubrics seem to have a positive effect on Performance/avoidance SRL, decreasing it more than do scripts.

Finally, Panadero and Romero (2014) compared the effect of a ‘baseline SA condition’ (i.e., asking the students to self-assess their work but without providing any instrument to facilitate the SA) to a ‘rubric referenced SA condition’ for pre-service teachers. The rubric group scored higher on Learning SRL and formulated more accurate SA, as compared to the baseline condition. However, the students in the rubric group also reported higher levels of stress while performing the task. The decrease in Performance/avoidance SRL was larger in the baseline SA condition, but the Performance/avoidance SRL scores of the rubric group decreased significantly as well.

Two main conclusions can be drawn from the work by Panadero and colleagues. First, promoting SA can increase the use of SRL strategies. Second, different instruments, such as rubrics and scripts, may have different effects on student SRL: scripts seem to increase Learning SRL, while rubrics may decrease Performance/avoidance SRL, which is often detrimental for learning.

### ***18.4.5 The Role of the Teacher in Promoting SA: Guidelines for Implementation***

The teacher’s role in facilitating students’ SA and in giving opportunities for practice has been emphasized in AfL research. In light of the large number of recommendations in the literature on implementation of SA in classrooms, we will highlight some guidelines that we consider to have special relevance for the development of SRL. Andrade and Valtcheva (2009) and Ross (2006) proposed several recommendations regarding instructional conditions to support SA. The following list combines their recommendations:

1. Define the criteria by which students assess their work
2. Teach students how to apply the criteria
3. Give students feedback on their self-assessments
4. Give students help in using self-assessment data to improve performance
5. Provide sufficient time for revision after self-assessment
6. Do not turn self-assessment into self-evaluation by counting it toward a grade.

The sixth recommendation concerns a complex and controversial issue, namely, the use of SA for summative purposes. On the one hand, students' assessment of their own work may be evaluated and taken into consideration by a teacher when grading student performance. This is done by a number of teachers (Panadero et al. 2014b), and some teachers (in higher education) consider SA to be a potentially more valid measure of students' achievement than their results on traditional exams (Tan 2012). On the other hand, using students' self-attributed grades for summative purposes may encourage students to make strategic choices in order to maximize their chances of achieving a higher grade, such as overestimating their performance (Boud and Falchikov 1989), instead of focusing on learning and improving their SA and SRL skills.

With respect to teachers' preparation as facilitators of SA, Panadero et al. (2014b) found that the strongest predictors for teachers' use of SA in their classroom were: (a) previous positive experience with SA, (b) endorsement of the educational advantages of SA (detection and correction of problems, saving time for the teacher, improvement of students' learning by using SA), and (c) previous training in assessment courses. These three aspects could be enhanced by means of 'Teacher Learning Communities' (TLCs), which have been suggested as a means to facilitate changes in the deep-rooted practices and habits of 'traditional assessment' (Wiliam and Thompson 2007). The use of professional learning communities for such purposes received further support by an overview of research on professional development and teacher learning by Borko (2004). She concluded that there is evidence suggesting that strong professional learning communities can not only foster teacher learning and instructional improvement but also contribute to school reform. Organizing such groups may therefore aid teachers in implementing new instructional practices, such as SA. It should be noted, however, that the development of teacher communities may be difficult and time consuming.

## 18.5 Peer Assessment

### 18.5.1 PA Definition

PA is 'an arrangement in which individuals consider the amount, level, value, worth, quality or success of the products or outcomes of learning of peers of similar status' (Topping 1998, p. 250). However, there is a large variety of PA practices that differ in terms of (a) purpose (summative vs. formative), (b) format (marking/rating with or without comments/feedback), and (c) degree of interaction between the assessor and assessee (e.g., PA of individual performance vs. PA of fellow group members' contribution to group work) (Strijbos et al. 2009).

### ***18.5.2 Relationship of PA and Co-Regulation***

The presence of the ‘other’ is emphasized in the literature devoted to the development of SRL skills and is deeply grounded in a Vygotskian perspective on learning and development (McCaslin 2009; McCaslin and Hickey 2001). According to this perspective, students develop their skills in a context where they can observe and emulate significant others, an event known as ‘co-regulation’ that ‘connotes shared responsibility’ (McCaslin and Hickey 2001, p. 243). In situations of co-regulation, the teacher’s assessment of students’ progress interacts with the students’ own attributions and interpretations:

Self-evaluation of personal progress is a central feature of social learning theory. In a Vygotskian tradition, self-evaluation is as much about personal meaning and affect as it is about progress toward standards, especially those set by others. (McCaslin and Hickey 2001, p. 248)

Furthermore, co-regulation through interaction with peers aligns well with the Vygotskian notion of the zone of proximal development (ZPD), which delineates what the student can do with some scaffolding and help from others. In PA, the peer acts as a source of such help and thus as a co-regulator of learning by the student who receives the PA. Naturally, this hinges on the quality of the PA provided as well as whether the assessee agrees with the PA and uses the suggestions.

Recent research on the ‘role of the other’ in SRL may be divided into two orientations: (1) ‘Co-regulated learning,’ which are situations where a temporary coordination of regulation occurs between the student and a significant other (i.e., teacher or peer), and (2) ‘Shared regulation of learning,’ where the regulatory processes are interdependent among the students who are participating in a collaborative task (Hadwin et al. 2011). In co-regulated learning, the student’s interaction with others allows the student to internalize regulatory processes. In socially shared regulation of learning, the students work as a coherent team to attain common goals. Furthermore, in socially shared regulation, all students participate equally in the regulation of each other’s actions, whereas in co-regulation, the student interacts with a person who has a superior or more expert role (teacher or more knowledgeable peer). Recent research has shown that both shared- and co-regulation can be empirically differentiated, with shared regulation enhancing group performance and the use of more advanced shared strategies (Panadero and Järvelä 2015).

In terms of co-regulation, teachers often play the part of the significant other via the assessments they produce by acting as a role model or by guiding and/or assessing the students who need to learn how to regulate their learning (Andrade and Brookhart 2014). As such, classroom assessment is conceptualized as a way to promote students’ regulation of learning, especially if the assessment practices follow the principles of AfL (Andrade and Brookhart 2014). In addition, teachers can provide opportunities for students to act as co-regulators of their peers’ learning via PA, which has simultaneously an impact on the peer assessors’ SRL skills (Nicol and McFarlane-Dick 2006).

### ***18.5.3 Empirical Evidence of the Relationship Between PA and Co-Regulation***

Empirical evidence of a connection between PA as an instance of co-regulation and enhanced outcomes is scarce. One example is the research by van Zundert (2012) and her colleagues. In a review of effective peer-assessment processes, van Zundert et al. (2010) found four studies showing that peer feedback (the main PA process) ‘positively influenced domain-specific skill’ (p. 274). It should be noted that none of the four studies explicitly positioned PA as an instance of co-regulation. The interaction between domain-specific skills and PA was further studied in two empirical investigations (van Zundert et al. 2012a, b). One of the conclusions from these studies is that it is probably better if students develop some of the domain-specific skills *before* they are asked to peer assess. This conclusion was drawn from a cognitive load theory standpoint: when learning novel tasks, students’ cognitive capabilities are directed towards the activity at hand, which means that there is not enough ‘cognitive space’ for handling additional demanding activities, such as PA (van Zundert et al. 2012b). This implies that for students to act as effective co-regulators, they need sufficient domain-specific knowledge and skills. The minimum required degree of domain-specific knowledge and skills is, however, an open issue and it may very well be domain and task dependent. In sum, despite the theoretical connections between PA and co-regulation, there is clearly a need for future research to explore this connection explicitly and in more detail.

### ***18.5.4 The Role of the Teacher in Promoting PA: Guidelines for Implementation***

While SA may occur without the teacher promoting such practice in the classroom, teacher intervention is almost always necessary to introduce formal PA as part of their classroom AfL practice. Another difference between SA and PA is that it might be even more important to consider the tensions between summative and formative uses of PA. For example, Panadero (2016), reviewed research on interpersonal factors of PA and concluded that formative approaches to PA (i.e., approaches including peer feedback and the possibility to interact with each other) seem to reduce the impact of negative interpersonal factors (e.g., feelings of unfairness), whereas summative approaches to PA (such as peer grading) can reinforce tensions among assessors and assessees. Nevertheless, it has been shown that some approaches to PA that combine formative and summative purposes can support student learning (depending on the quality of the feedback) and that tensions can be mitigated if the PA score is explained via peer feedback (Panadero 2016). In summary, teachers need to be aware of both the benefits and the limitations of PA in order to make informed decisions on how to implement PA in their classrooms. In the worst-case scenario, formative assessment intentions might be

transformed into recurrent and fragmented summative assessments, which may impair the solidarity of the students in the class.

With these considerations in mind, we compiled some guidelines for teachers' implementation of PA. These guidelines are based on a list of principles proposed by Topping (2003), but we removed recommendations referring to strictly summative approaches to PA and rephrased some recommendations in such a way that they apply to various types of PA. When following these recommendations, teachers will still need to tailor proposed PA practices in relation to the learning goals, the task, and the specific context of each individual classroom.

1. Clarify the purpose of PA, its rationale and expectations to the students
2. Involve students in developing and clarifying assessment criteria
3. Match participants (e.g., individuals, groups) in a way that fosters productive PA
4. Determine the PA format (e.g., rating with or without comments) and mode of PA interaction (e.g., face-to-face or online)
5. Provide quality PA training, examples and practice (including feedback about PA)
6. Provide rubrics, scripts, checklists, or other tangible scaffolding for PA
7. Specify PA activities and timescale
8. Monitor the PA process and coach students.

## 18.6 General Conclusions

AfL practices refer to a type of educational assessment in which student involvement and feedback are central. In other words, assessment is used to generate feedback by and for the students. Two of the most important AfL practices in this respect are SA and PA. In this chapter we discussed the connections of both SA and PA to SRL. It is evident from this discussion that there is both theoretical and empirical support for the relationship between SA and SRL and that training in SA may indeed enhance the use of SRL skills. In contrast, the relationship between PA and co-regulation is still very much implicit; very few theoretical and empirical studies have investigated this issue.

Even if the promotion of SA has been shown to increase the use of SRL in different contexts, it is currently not possible to identify which specific interventions or instruments would be the most effective in enhancing SRL skills. However, the formulation of explicit criteria for SA purposes, so that the students may set realistic goals and evaluate their progress (both during the process and afterwards), seems to be a particularly promising way forward. As has been shown, explicit criteria may support all of the SRL phases and several recommendations for the implementation of SA in the classroom are connected to the use of such criteria. For example, teachers should make the criteria by which students assess their work explicit, teach the students how to apply the criteria, give students feedback on their SA, help

students in using SA information to improve their performance, and provide sufficient time for revision after SA.

Although there is less empirical evidence that directly supports the relation between SRL and PA, the available evidence suggest that it is of utmost importance to clarify the rationale for PA in advance, involve students in determining the criteria, clearly specify the PA format as well as how students are supposed to interact, and provide them with sufficient training and scaffolds to conduct the PA activities. Scaffolding may be particularly important in situations where students' domain-specific knowledge and skills are limited.

Nevertheless, a major challenge for the implementation of both SA and PA is that many teachers—when given the choice—prefer to not promote students' active involvement in assessment (e.g., Jonsson et al. 2015; Panadero and Brown 2015). It is thus important to develop recommendations based on the research literature about how to support teachers in their implementation of SA and PA, as part of an overall AfL approach to classroom assessment, for instance, by encouraging teachers to work together in professional learning communities. In such communities, teachers may be encouraged by colleagues to work with SA and/or PA, which in turn may affect whether the teachers are likely to engage with these pedagogical resources in their classrooms. Furthermore, such communities may also facilitate discussions on effective designs for SA and PA, and on how teachers can best implement them to foster student learning and SRL skills. After all, having positive experience with SA is one of the strongest predictors for further use of SA (Panadero et al. 2014b).

Finally, the implementation of both SA and PA entails risks: (a) a formative assessment activity intended to support student learning could turn into a solely summative event, and (b) if poorly designed, SA or PA could become an activity in itself that consumes valuable classroom time without necessarily contributing effectively to student learning. Although the use of SA and PA has been shown to have several positive effects on student motivation and learning, there are no well-designed studies showing positive effects of self- or peer grading. Summative uses of SA and PA seem to have a number of negative effects which could counteract the intentions of AfL, such as students overestimating their own performance or providing less than constructive feedback to particular peers. Nevertheless, given these caveats, an increased research effort in exploring the effects of AfL practices—and of SA and PA in particular—constitutes a promising classroom assessment perspective, especially if recommendations such as the ones provided here, are implemented.

**Acknowledgements** This chapter is based in part on research funded by the Spanish Ministry (Ministerio de Economía y Competitividad) via Ramón y Cajal funding granted to the first author (File id. RYC-2013-13469).

## References

- Andrade, H., & Brookhart, S. M. (2014). *Assessment as the regulation of learning*. Paper presented at the American Educational Research Association, Philadelphia, USA.
- Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation*, 10(3), 1–11. <http://pareonline.net/getvn.asp?v=10&n=3>
- Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Into Practice*, 48(1), 12–19.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, 5(1), 7–73.
- Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology*, 54(2), 199–231.
- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, 33(8), 3–15.
- Boud, D., & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher-education: A critical analysis of findings. *Higher Education*, 18(5), 529–549. doi: [10.1007/BF00138746](https://doi.org/10.1007/BF00138746)
- Brown, G. T. L., & Harris, L. R. (2013). Student self-assessment. In J. McMillan (Ed.), *The SAGE handbook of research on classroom assessment* (pp. 367–393). Thousand Oaks, CA: SAGE.
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65(3), 245–281.
- Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, 24(2), 205–249. doi:[10.1007/s10648-011-9191-6](https://doi.org/10.1007/s10648-011-9191-6)
- de la Harpe, B., & Radloff, A. (2000). Informed teachers and learners: The importance of assessing the characteristics needed for lifelong learning. *Studies in Continuing Education*, 22(2), 169–182. doi:[10.1080/01580379950177323](https://doi.org/10.1080/01580379950177323)
- Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, 3, 231–264. doi:[10.1007/s11409-008-9029-x](https://doi.org/10.1007/s11409-008-9029-x)
- Hadwin, A. F., Järvelä, S., & Miller, M. (2011). Self-regulated, co-regulated, and socially shared regulation of learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 65–84). New York: Routledge.
- Harris, L. R., & Brown, G. T. L. (2013). Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: Case studies into teachers' implementation. *Teaching and Teacher Education*, 36(0), 101–111. doi:[10.1016/j.tate.2013.07.008](https://doi.org/10.1016/j.tate.2013.07.008)
- Jonsson, A. (2013). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education*, 14(1), 63–76. doi:[10.1177/1469787412467125](https://doi.org/10.1177/1469787412467125)
- Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment & Evaluation in Higher Education*, 39(7), 840–852. doi:[10.1080/02602938.2013.875117](https://doi.org/10.1080/02602938.2013.875117)
- Jonsson, A., Lundahl, C., & Holmgren, A. (2015). Evaluating a large-scale implementation of assessment for learning in Sweden. *Assessment in Education: Principles, Policy & Practice*, 22(1), 104–121. doi:[10.1080/0969594X.2014.970612](https://doi.org/10.1080/0969594X.2014.970612)
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 130–144. doi:[10.1016/j.edurev.2007.05.002](https://doi.org/10.1016/j.edurev.2007.05.002)
- Kostons, D., van Gog, T., & Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learning and Instruction*, 22(2), 121–132. doi:[10.1016/j.learninstruc.2011.08.004](https://doi.org/10.1016/j.learninstruc.2011.08.004)
- Kramarski, B., & Michalsky, T. (2009). Three metacognitive approaches to training pre-service teachers in different learning phases of technological pedagogical content knowledge. *Educational Research and Evaluation: An International Journal on Theory and Practice*, 15(5), 465–485.



- Lan, W. Y. (1998). Teaching self-monitoring skills in statistics. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning: From teaching to self-reflective practice* (pp. 86–105). New York: Guilford Press.
- McCaslin, M. (2009). Co-regulation of student motivation and emergent identity. *Educational Psychologist*, 44(2), 137–146. doi:[10.1080/00461520902832384](https://doi.org/10.1080/00461520902832384)
- McCaslin, M., & Hickey, D. T. (2001). Self-regulated learning and academic achievement: A Vygotskian view. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement* (pp. 227–252). New York: Lawrence Erlbaum Associates.
- Nicol, D., & McFarlane-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. doi:[10.1080/03075070600572090](https://doi.org/10.1080/03075070600572090)
- Olina, Z., & Sullivan, H. J. (2004). Student self-evaluation, teacher evaluation, and learner performance. *Educational Technology Research and Development*, 52(3), 5–22. doi:[10.1007/BF02504672](https://doi.org/10.1007/BF02504672)
- Panadero, E. (2011). *Instructional help for self-assessment and self-regulation: Evaluation of the efficacy of self-assessment scripts vs. rubrics*. Doctoral dissertation, Universidad Autónoma de Madrid, Spain.
- Panadero, E. (2016). Social, interpersonal and human effects of peer assessment: A review and future directions. In G. T. L. Brown & L. R. Harris (Eds.), *Human factors and social conditions of assessment* (pp. 247–266). New York: Routledge.
- Panadero, E., & Alonso-Tapia, J. (2013). Self-assessment: Theoretical and practical connotations. When it happens, how is it acquired and what to do to develop it in our students. *Electronic Journal of Research in Educational Psychology*, 11(2), 551–576. doi:[10.14204/ejrep.30.12200](https://doi.org/10.14204/ejrep.30.12200)
- Panadero, E., & Alonso-Tapia, J. (2014). How do students self-regulate? Review of Zimmerman's cyclical model of self-regulated learning. *Anales De Psicología*, 30(2), 450–462. doi:[10.6018/analesps.30.2.167221](https://doi.org/10.6018/analesps.30.2.167221)
- Panadero, E., Alonso-Tapia, J., & Huertas, J. A. (2012). Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and Individual Differences*, 22(6), 806–813. doi:[10.1016/j.lindif.2012.04.007](https://doi.org/10.1016/j.lindif.2012.04.007)
- Panadero, E., Alonso-Tapia, J., & Huertas, J. A. (2014a). Rubrics vs. self-assessment scripts: Effects on first year university students' self-regulation and performance. *Infancia y Aprendizaje: Journal for the Study of Education and Development*, 37(1), 149–183.
- Panadero, E., Alonso-Tapia, J., & Reche, E. (2013). Rubrics vs. self-assessment scripts effect on self-regulation, performance and self-efficacy in pre-service teachers. *Studies In Educational Evaluation*, 39(3), 125–132. doi:[10.1016/j.stueduc.2013.04.001](https://doi.org/10.1016/j.stueduc.2013.04.001)
- Panadero, E., & Brown, G. T. L. (2015). *Higher education teachers' assessment practices: Formative espoused but not yet fully implemented*. Paper presented at the Fifth Assessment in Higher Education Conference 2015, Birmingham, (UK).
- Panadero, E., Brown, G. T. L., & Courtney, M. G. R. (2014b). Teachers' reasons for using self-assessment: A survey self-report of Spanish teachers. *Assessment in Education: Principles, Policy & Practice*, 21(3), 365–383. doi:[10.1080/0969594X.2014.919247](https://doi.org/10.1080/0969594X.2014.919247)
- Panadero, E., Brown, G. T. L., & Strijbos, J. W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*. doi:[10.1007/s10648-015-9350-2](https://doi.org/10.1007/s10648-015-9350-2)
- Panadero, E., & Järvelä, S. (2015). Socially shared regulation of learning: A review. *European Psychologist*, 20(3), 190–203. doi:[10.1027/1016-9040/a000226](https://doi.org/10.1027/1016-9040/a000226)
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129–144. doi:[10.1016/j.edurev.2013.01.002](https://doi.org/10.1016/j.edurev.2013.01.002)
- Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice*, 21(2), 133–148. doi:[10.1080/0969594X.2013.877872](https://doi.org/10.1080/0969594X.2013.877872)
- Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, 36(2), 89–101.

- Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment Research & Evaluation*, 11. <http://pareonline.net/getvn.asp?v=11&n=10>. Accessed May 7, 2015.
- Segers, M., Nijhuis, J., & Gijssels, W. (2006). Redesigning a learning and assessment environment: The influence on students' perceptions of assessment demands and their learning strategies. *Studies in Educational Evaluation*, 32(3), 223–242. doi:10.1016/j.stueduc.2006.08.004
- Strijbos, J. W., Ochoa, T. A., Sluijsmans, D. M. A., Segers, M., & Tillema, H. H. (2009). Fostering interactivity through formative peer assessment in (web-based) collaborative learning environments. In C. Mourlas, N. Tsianos, & P. Germanakos (Eds.), *Cognitive and Emotional Processes in Web-Based Education: Integrating Human Factors and Personalization* (pp. 375–395). Hersey, PA: IGI Global.
- Tan, K. H. K. (2012). *Student self-assessment. Assessment, learning and empowerment*. Singapore: Research Publishing.
- Topping, K. J. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249–276.
- Topping, K. J. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimising new modes of assessment: In search of qualities and standards* (Vol. 1, pp. 55–87). Dordrecht: Springer.
- van Zundert, M. (2012). *Conditions of peer assessment for complex learning* (Doctoral dissertation, Maastricht University).
- van Zundert, M. J., Könings, K. D., Sluijsmans, D. M. A., & van Merriënboer, J. J. G. (2012a). Teaching domain-specific skills before peer assessment skills is superior to teaching them simultaneously. *Educational Studies*, 38(5), 541–557. doi:10.1080/03055698.2012.654920
- van Zundert, M. J., Sluijsmans, D. M. A., Könings, K. D., & van Merriënboer, J. J. G. (2012b). The differential effects of task complexity on domain-specific and peer assessment skills. *Educational Psychology*, 32(1), 127–145. doi:10.1080/01443410.2011.626122
- van Zundert, M., Sluijsmans, D., & Van Merriënboer, J. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction*, 20(4), 270–279. doi:10.1016/j.learninstruc.2009.08.004
- Wiliam, D. (2014). *Formative assessment and contingency in the regulation of learning processes*. Paper presented at the American Educational Research Association, Philadelphia, USA.
- Wiliam, D., & Thompson, M. (2007). Integrating assessment with learning: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53–82). Mahwah, NJ: Lawrence Erlbaum Associates.
- Winne, P. H. (2014). *Fusing self-regulated learning with formative evaluation to improve education*. Paper presented at the American Educational Research Association, Philadelphia, USA.
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–40). San Diego, CA: Academic Press.
- Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist*, 48(3), 135–147. doi:10.1080/00461520.2013.794676
- Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 299–315). New York: Routledge.
- Zimmerman, B. J., & Schunk, D. H. (2001). *Self-regulated learning and academic achievement*. New York: Lawrence Erlbaum Associates.
- Zimmerman, B. J., & Schunk, D. H. (2011). *Handbook of self-regulation of learning and performance*. New York: Routledge.

# Chapter 19

## Assessment for Learning: Co-Regulation *in and as Student–Teacher Interaction*

Margaret Heritage

**Abstract** Assessment for learning (AfL) conceptualizes assessment as integral to teaching and learning. Interaction is considered a primary source of evidence in AfL. This chapter examines sequences of one-on-one teacher–student interaction from transcribed classroom videos as instances of co-regulation. Co-regulation refers to the joint influence of student self-regulation, and of regulation from other sources, including teachers, on student learning. The notion of learning as co-regulation also accords with the perspective of situated cognition in which learners actively construct knowledge with others. Challenges to the implementation of interaction in AfL and the implications for co-regulated learning are discussed, as well as ways in which teacher practice in this area could be supported.

### 19.1 Introduction

Assessment for learning (AfL) conceptualizes assessment as integral to teaching and learning. AfL occurs in the flow of activity and transactions in the classroom, and has as its central focus pedagogical intervention in the immediacy of student learning and the students' agency in the learning process (Swaffield 2011). The practice of AfL involves establishing clear learning goals and performance criteria (success criteria), eliciting and interpreting evidence of learning against the criteria while that learning is developing, promoting student agency, and making immediate or near-immediate adjustments to teaching and learning based on evidence (Black et al. 2003; Black and Wiliam 1998, 2009; Sadler 1989).

The origins of the term assessment are to be found in the Latin verb *assidere*, to sit beside, and indeed, interaction between teachers and students has been characterized as a principal source of evidence in AfL (Allal 2010; Black and Wiliam 2009; Heritage 2013a, b; Torrance and Pryor 1998). Questions designed to explore students' thinking (Black and Wiliam 2005; Shavelson et al. 2008), assessment

---

M. Heritage (✉)  
WestEd, San Francisco, USA  
e-mail: mherita@wested.org

conversations conceived of as dialogue embedded in ongoing classroom activity (Ruiz-Primo and Furtak 2006), and one-on-one interactions between a student and teacher (Heritage and Heritage 2013) are examples of obtaining evidence from which to draw inferences about students' current learning status.

This chapter expands the idea of assessment as 'sitting beside' and examines sequences of one-on-one teacher–student interaction when a teacher and student sit together as occasions for AfL. In it, I develop the idea of interaction as a source of evidence in AfL by specifically considering these one-on-one teacher–student interactions as instances of co-regulation. Co-regulation refers to 'the joint influence of student self-regulation, and of regulation from other sources—teachers, peers, curriculum materials, assessment—on student learning' (Allal 2011, p. 332). This view of learning accords with the perspective of situated cognition in which learners actively construct knowledge with others. The focus is shifted from the behavior and cognition of individuals (e.g., as in the cognitive and behavioral perspectives) to a focus on cognitive agents interacting with each other. This situative approach treats knowledge as public and as distributed among individuals and their environment, including the objects, artifacts, tools, and communities that comprise the learning context (Greeno 2006; Greeno et al. 1996).

In a key observation, Bandura (1991, p. 249) notes: 'If human behavior were regulated solely by external outcomes, people would behave like weathervanes, constantly shifting direction to conform to whatever momentary social influences happen to impinge on them.' In contrast, he observes that human behavior instead is regulated by individuals through the self-monitoring of behavior, its causes and effects; judgment of behavior in relation to standards and circumstances; and reflection upon actions. In the realm of learning, these mechanisms of human agency—self-regulation—can be conceptualized as the processes by which learners set goals, execute, reflect on, and adapt learning in order to optimize learning (Hadwin et al. 2011).

Self-regulation is conducted by individuals as a private process, whereas co-regulation involves shifts in the management of regulation between the learner and another, with assistance for self-regulation provided by external modeling and feedback (Järvelä and Järvenoja 2011). Noting that research about co-regulation focuses on interactions between individuals and others, Hadwin et al. (2011) characterize co-regulated learning as emergent interaction providing temporary support in the service of self-regulation. In the next section, I examine two sequences of teacher–student interaction and the pedagogical mechanisms employed in co-regulation.

## 19.2 Teacher–Student Interactions

Two hours of videotaped interaction were collected from a third-grade mathematics classroom (students aged 8–9) and a fifth-grade writing classroom (students aged 10–11) in an elementary school located in the downtown area of a large American

city. All students were recipients of either free or reduced-price lunch, indicating the low socioeconomic status of their families. There were thirty students in the mathematics class, most of whom were classified as English-language learners, a designation which entitles the students to additional language support to attain proficiency in English. In the writing class, there were 26 students, and over 50 % were classified as English learners.

The videotapes were transcribed and examined for sequences of teacher–student interaction in one-on-one formative assessment. Specifically, the interaction was examined from the perspective of the following pedagogical mechanisms drawn from Heritage’s (2010) discussion of question design:

1. Agenda setting: how questions set topics for response and the kinds of responsive actions that are established from respondents;
2. Presupposition: what fundamental assumptions are being made in the delivery of questions;
3. Epistemic gradient (Heritage 2013a, b): the gap in knowledge between questioner and the respondent is encoded by the grammatical form of the question: interrogative syntax (e.g., ‘are you focusing on grammar?’) reflects a relatively greater gap in knowledge between question and answer, that is, deeper epistemic gradient; declarative syntax (e.g., ‘you’re focusing on grammar?’) reflects a smaller gap in knowledge between speakers, that is, relatively shallow epistemic gradient;
4. Preference: involves primarily polar questions—or yes/no questions—and concerns whether the question is tilted toward a yes/no response and the extent to which that tilt is built into the question design. For example, the question ‘So is there a way that you think maybe you can combine those two into one?’ is tilted towards a ‘yes’ answer, whereas ‘There isn’t a way that you can combine those two into one is there?’ favors a ‘no’ answer.

Student–teacher dialogues were examined for the ways in which they addressed these characteristics.

### ***19.2.1 Third-Grade Mathematics Classroom***

The first sequence of interaction is from a third-grade mathematics classroom. The students are working on writing mathematical expressions with parentheses and are given a word problem on the board, which they are to solve using expressions and parentheses:

Expressions with parentheses

Nia earned \$11 for pulling weeds and \$10 for cleaning windows. She spent \$7 on a movie ticket and \$3 for snacks. She saved the rest of the money. How much money did Nia save?

The teacher, Ms. Castro, preconstructs the task in terms of ‘three steps’ to solve the problem in the following sequence, which eventuates in the class stating that they need to identify the question, the clues, and the key words:

- 1 Ms. C: I want you to discuss what do we need to know in order to solve  
 2 this problem? And there are three steps. Who can remember one  
 3 of the three steps to solve this problem? Daniel?  
 4 ST1: Identify a question.  
 5 Ms. C: Okay. Then we identify a question.  
 6 ST2: We identify the clues.  
 7 Ms. C: Okay. Then we identify the clues. Mm hm, go ahead.  
 8 ST3: Look for key words.  
 9 Ms. C: And key words. Okay. I want you all to turn to your partner and  
 10 I want you to identify all three. The questions, the clues,  
 11 remember there might be more than just one, and the key words.  
 12 Now go ahead and discuss that together.

In this classroom, learning routinely takes place in the context of joint activity where students collaborate to solve problems, activating each other’s knowledge and discussing possible solutions, so the students move quickly into pairs and begin to discuss the clues in the problem. While the students are discussing the problem, Ms. Castro moves around the pairs and explores their progress. Approximately eight minutes into the partner discussion, Ms. Castro reaches Rico and his partner. Rico has used different color highlights to identify the question, key words, and clues to the problem in his exercise book. He has also made a representation of how he solved the problem using expressions with parentheses. Ms. Castro sees that he has correctly solved the problem with his representation but wants to explore the reasoning that underlies his result.

Ms. Castro begins with a question designed to identify Rico’s initial approach to the problem and establishes that he has used the steps outlined at the beginning of the lesson (lines 1–10). She determines that he has identified several key words needed to create mathematical expressions to solve the problem.

- 1 Ms. C: And what was your first step in solving this problem?  
 2 Rico: First is to underline questions, clues, and key words.  
 3 Ms. C: Okay.  
 4 Rico: (pointing at paper) These are the questions and the key  
 5 words and the clues.  
 6 Ms. C: So, can you go over with me some of the key words you  
 7 found in this problem?  
 8 Rico: I found that she earned, spent, saved, and ( ).  
 9 Ms. C: So there were several key words. There wasn’t just one key  
 10 word in particular, right?

Ms. Castro’s question at lines 6–7 clearly presupposes the existence of more than one key word in the problem, and Rico’s response, in listing at least three, confirms

that. Ms. Castro's subsequent response at lines 9–10, underscores Rico's response, summarizing it and presenting it (with 'so') as a conclusion drawn from Rico's experience. At this point, she turns to Rico's numerical representation of the problem and asks him to explain the thinking behind this representation—in essence, how the language of the problem was translated into a mathematical expression.

Ms. Castro's slightly awkward question (lines 15–16 below)—'Can you show me how you came to that conclusion of using that expression?'—is designed to avert a focus on the answer, in favor of how he arrived at the representation, which is her target of interest.

- 12 Ms. C: Now I noticed that you went a step further and you tried  
 13 to solve this problem. (points at paper) Now I also  
 14 noticed that you used an expression rather than  
 15 an equation. Can you— How did you do this? Can you show me  
 16 how you came to that conclusion of using that expression?

This question invites an extended explanation of his thinking, in which Rico shows that he first adds items labeled 'earned' and also items labeled 'spent' before going on to subtract the latter from the former. He also explains that he used the parentheses to remember which items fell in which categories.

- 17 Rico: Well like she earned 11 for pulling weeds and 10 for cleaning  
 18 windows so if she earned them so I added them because it said  
 19 that she spent 7 dollars on a movie ticket and 3 for a snack,  
 20 so I used the math to do this because then I get confused,  
 21 So I— Instead I added the ones that she earned and added the  
 22 ones that she spent it on, (0.3) and just to show that— just  
 23 so I won't get mixed up I put parentheses because to learn  
 24 that uhm ( ) that I am going to subtract them.  
 25 (2.0)

- 26 Ms. C: So I see— So you put parentheses around the 11 and the 10,  
 27 then you put parentheses around 7 plus 3. (0.2) Right?

- 28 Rico: Yes.

Once again, Ms. Castro (lines 26–27) offers a succinct understanding of the reasoning described in Rico's previous turn, while referencing and focusing on the specific mathematical representation ('parentheses') that Rico mentioned at line 23. In the next segment, Ms. Castro explores Rico's next step after putting the numbers in the parentheses, focusing on his understanding of the need to solve the operation expressed in each parenthesis prior to any subsequent step. Finding Rico's understanding to be solid verbally (as well as representationally), she proceeds to his final step in the problem—subtracting the money spent from the money earned.

- 29 Ms. C: Now what's your next step after putting parenthesis around  
 30 ( )?  
 31 Rico: To add what's in thuh— in thuh open parenthesis and the

- 32 closed, what's inside of it. [I have to add it.
- 33 Ms. C: [So:
- 34 Ms. C: Your first step was going to be to solve what's inside the  
 35 parenthesis. Okay. Why don't you go ahead and show us  
 36 what you came up with.
- 37 Rico: I came up with 21 with the 11 plus 10, and 10 is the 7 plus  
 38 3, and then since I figured out that the minus was still there,  
 39 so I put it there and I did 21 minus 10, which gave me 11.
- 40 Ms. C: And eleven: (0.2) Is that your final answer?
- 41 Rico: Eleven dollars. (Adds \$ sign to his result)

In this final segment, Ms. Castro uses an open-ended question that invites Rico to describe the next step in his thinking. After he has done so (lines 31–32), she briefly summarizes his response, in the process replacing his use of the word ‘add’ (line 32) with the more inclusive technical mathematical expression ‘solve’ (line 34). She then projects a shift to a new element in the exposition with the word ‘okay’ (Beach 1993), and invites Rico to present his conclusion. It may be noted here that Ms. Castro’s turn at lines 34–36, by treating the word ‘solve’ as unproblematic, fully assumes that he will grasp its meaning.

Before this interaction, Ms. Castro had already obtained two sources of evidence from Rico’s notebook: his highlighted elements of the clues and key words, and his representation of a solution to the problem. While she could have taken both elements as evidence of understanding, she uses her interaction as an opportunity to probe his thinking more deeply. She wants to be sure she has sufficient information to make a judgment about his learning status (Smith 2003). Ms. Castro probes Rico’s thinking through open-ended questions (lines 1, 15–16, and 29) that permit Rico to establish the sequencing of his thinking and, at each next step, enable him to determine the way forward. The result is that her questioning, though sequentialized in terms of the problem-solving pattern, is one that allows the nature of Rico’s thinking to emerge with only minimal scaffolding. In this AfL process, Ms. Castro can become fully assured that Rico grasps the whole problem, including translating the language of the question into the language of mathematics and the logical series of steps required to solve the problem. At the conclusion of this segment, Ms. Castro tests Rico’s commitment to his final answer to the problem (line 40). Rico’s confirmatory response—accomplished via repetition, rather than a simple ‘yes’ (Raymond 2003; Heritage and Raymond 2012)—assumes complete ownership of the solution, and he underscores this ownership by simultaneously adding a dollar sign to his circled written answer.

### ***19.2.2 Fifth-Grade Writing Classroom***

This example focuses on the social interactional context of the pedagogical mechanisms and the way in which teacher and student respect one another as



co-regulators of the student's experience. Accordingly, in this example, I focus on four basic elements in the exchange:

1. the student's initiative soliciting feedback from her teacher;
2. the establishment of a collaborative relationship between teacher and student;
3. the agentive role maintained by the student throughout the interaction; and
4. the suggestive quality of the feedback.

In Ms. Lara's writing class, the students are learning about argument structure. Before the sequence of interaction below, the students had learned about arguments and counterarguments and are now using these structures in their own writing. Ms. Lara's instruction occurs within the predictable routine of a 'writer's workshop' setting (Calkins 1994). Each session of the workshop begins with a mini-lesson focused on argument structure, which is followed by a period in which the students engage in independent writing, using what they learned in the mini-lesson to further their work, and soliciting feedback from peers as their writing develops.

Angie is involved in independent writing when Ms. Lara comes to sit beside her and engages in the following conversation:

- 1 Ms. L: Ok Angie, what are you working on?  
 2 Angie: I'm working on my final draft, and wanted to make it kind of  
 3 sentences, and I wanted your feedback.  
 4 Ms. L: Okay. Do we have our success criteria here, our checklist?  
 5 Angie: Yes.  
 6 Ms. L: What are you looking at right now, what are you focusing on? Are you  
 7 focusing on punctuation? Are you focusing on grammar?  
 8 Angie: I'm working on this one.  
 9 Ms. L: Oh clarity, so you're asking yourself if this is going to make sense  
 10 to somebody who had no idea. So what do you think so far?  
 11 Angie: I don't know if I should, because I started with two questions and  
 12 then I ended with a period. And then I started another question.  
 13 Ms. L: I see, so let's read it and see how that makes sense.  
 14 Angie: It says 'The world has been taken by trash... What are you going to do  
 15 to save our earth?' [Angie continues to read her writing]  
 16 Ms. L: Ok, lets go back to your original concern. So you're concerned about  
 17 having two questions at the beginning. Well, the question that you  
 18 have here at the beginning 'I wonder why people don't pick up the  
 19 trash up?' Well following that up with what, what is this? 'People may  
 20 argue that,' what is that?  
 21 Angie: That's a counter argument.  
 22 Ms. L: That's a counter argument. So this question, 'I wonder why people  
 23 don't pick up trash?'  
 24 Angie: Is connected to my counterargument.  
 25 Ms. L: Is connected to your counterargument. So it makes sense. Okay? So  
 26 what's the other question that you feel maybe...  
 27 Angie: I was going to put, right here after about 3 billion people don't

- 28 care about the earth. I was going to put, I wonder why they don't  
 29 care. And then I was going to put this one.
- 30 Ms. L: Oh I see.
- 31 Angie: And I wanted to know if that was okay. To put two questions in a  
 32 question, period, and another question.
- 33 Ms. L: Well I think that 'I wonder why they don't care' and 'I wonder why  
 34 people don't pick up trash,' it's connected. It's connected. So is  
 35 there a way that you think maybe you can combine those two into one?  
 36 So that you don't have two questions back to back?
- 37 Angie: Yeah.
- 38 Ms. L: So can you think about that? Because 'I wonder why they don't care'  
 39 and 'I wonder why people don't pick up trash'
- 40 Angie: Are the same.
- 41 Ms. L: Are connected to each other, so you can definitely think about  
 42 connecting those two so that it's one question. But that has those two  
 43 things; those two components that you wanted to make sure that were  
 44 in there.
- 45 Angie: Okay.
- 46 Ms. L: Okay, so go ahead and think about how you can do that.

### **19.2.2.1 The Student's Initiative in Soliciting Feedback from Her Teacher**

First, it is apparent that Angie is soliciting assistance in her writing and she does so in response to Ms. Lara's question: 'what are you working on?' In going beyond the agenda of Ms. Lara's question (in lines 2–3), Angie clearly takes the initiative in requesting assistance. Second, her formulation of this help as 'feedback' invokes a relationship with Ms. Lara and displays a settled understanding of the routines of teaching and learning in Ms. Lara's writing class. Angie has already written a rhetorical question, 'I wonder why people don't pick up the trash?' and she is entertaining the idea of a second question back to back: 'I wonder why people don't care [about the earth]?'

### **19.2.2.2 The Agentive Role of the Student in a Collaborative Relationship Between Teacher and Student**

It is not difficult to identify passages in this exchange in which Angie takes an agentive discursive role. In addition to her initiation of the topic of the conversation at lines 2 and 3, and at lines 11 and 12, Angie articulates her problem as 'starting with two questions.' Moreover, at lines 24 and 40, Angie finishes the teacher's sentence, in effect completing her thought process for her. In short, Angie's conduct

indexes her view of the transaction as one involving a collaboration in which she has equal rights to initiate and pursue exchanges.

While Angie's conduct is presumptive of a collaborative relationship with her teachers, this presumption would be of little value if it were not reciprocated by the teacher. In this vein, at lines 4 and 6–7, Ms. Lara invites Angie to set the agenda for the conversation, which Angie duly does at line 8. At lines 16–17, Ms. Lara checks her understanding of Angie's problem—'So you're concerned...'—and finds by looking at Angie's text that her rhetorical question is immediately followed by a passage that is clearly identifiable as the beginning of a counter argument, lines 19–25. Having arrived at this point, Ms. Lara is ready to engage Angie's question about putting two questions consecutively: a concern that Angie specifically renews at lines 31–32. Without addressing this question directly, Ms. Lara observes that the two questions are closely connected and asks if Angie can think of a way of combining 'those two into one' (line 35). Across this entire sequence, Ms. Lara sustains an interactional stance in which her job is to come to an understanding of the problem with which Angie is presently grappling.

### **19.2.2.3 The Suggestive Quality of the Feedback**

From line 35 onwards, Ms. Lara offers Angie a way out of her dilemma: combining the two questions into one 'so that you don't have two questions back to back' (line 35–36). Subsequently, she invites Angie to think about how she can connect the two questions and to 'go ahead and think about how you can do that.' While she addresses the question issue as Angie's problem, she does not endorse Angie's position as her own, thus leaving the possibility open for Angie to decide to keep two questions back to back. Importantly, moreover, she offers Angie a hint rather than a concrete suggestion about how to resolve her problem. In sum, Ms. Lara, avoids making an authoritative determination as to whether Angie's concern is well founded or not and leaves her to find her own way to a solution. In these ways, she sustains Angie's role as an agentive writer and with it the collaborative relationship between the two that was established at the outset of the interaction.

### **19.2.2.4 Intersubjectivity and Co-Regulation**

A striking feature of Ms. Lara's interaction with Angie is the extent to which, while inviting Angie to take the lead in developing the topic of their conference, she is also careful to spell out her understandings of Angie's concerns (lines 9–10, 16–17) and to confirm Angie's interventions, such as the one at line 24 which completes her own preceding utterance (at lines 22–23). These actions are the small stitches that weave together an interaction in which contributions are both intersubjective and collaborative, indexing a relationship that is as much collegial as it is pedagogical. This process of interaction is a co-regulative one in which Angie's original concern is articulated, clarified, and worked up through a series of interchanges in

which additional considerations bearing on a possible solution to her problem are brought to the conversational surface and elaborated on.

### 19.3 Interaction as Co-Regulation

Co-regulation is grounded in intersubjectivity and scaffolding which involves participants in shared activity around a task (Järvelä and Järvenoja 2011). In the examples of teacher–student interaction presented above, each one had a particular focus where the student was already engaged in a learning task. In the mathematics example, the teacher wished to obtain evidence of Rico’s underlying understanding of the solution he had represented for the problem. In the writing class, the student’s concern about two back-to-back questions was the focus of the interaction. Intersubjective understanding between teacher and students was a central component in this process (Duranti 2010; Heritage 1984). It was made possible through the ways each successive turn in the interaction built upon the previous one, displaying an understanding of what was intended, and moving the sequence and its associated tasks forward (Schegloff 1992). As a result, the interactions of each dyad were conceptually and practically coordinated around specific tasks.

The conversation between Ms. Castro and Rico represents a paradigmatic form of co-regulation and does so at several levels. At the level of intersubjective sense making, as previously observed, both teacher and learner were active contributors to the management of understanding and its incremental extension over the course of the interaction. At a motivational level, the co-regulation was actively pursued in a spirit of collaboration, in which the two partners treated one another respectfully and as equals in the exchanges. In the process, Ms. Castro’s assurance of Rico’s mathematical understanding was secured and became intersubjectively ratified, while Rico’s ownership of his solution was consistently upheld and indeed underscored by Rico himself when he added the dollar sign to his final solution. At a third, organizational level, the management of classroom practice, routines, and expectations created an environment in which students were assumed to be independent learners and permitted Ms. Castro and Rico to have their exchange without interruption or disturbance. In terms of AfL, both participants have the information they need to make decisions about Rico’s next steps in learning—he can advance from these types of problems. While in this instance, Ms. Castro managed the regulation, Rico is enculturated into the process of self-regulation as he responds to Ms. Castro’s questions, bringing his thinking to a conscious level and making it possible for him to make a judgment about goal attainment (cf. Hadwin et al. 2011).

In the interaction between Angie and Ms. Lara, the management of regulation shifted between the two during the sequence. Angie initiated the interaction about her issue with two questions, evidencing her own regulatory process. Ms. Lara then managed the regulation throughout the subsequent turns, finally returning the management to Angie with her statement ‘so go ahead and think about how you can do that.’ From an AfL perspective, Angie’s next steps in moving forward with her

essay were clear to both her and her teacher. Angie's subsequent writing goal was confirmed intersubjectively as a result of the exchange about her concern and its possible resolution.

The meshing of understandings between the participants in each of the two interactional sequences is matched by a coordination of actions in which the questions and responses were reciprocally contingent upon each other. This is a hallmark of conversational interaction more generally (Sacks 1987; Schegloff 2007), but in the conversations that make up AfL sequences, this aspect of interaction is particularly significant. For example, in the math discussion, Ms. Castro asked Rico if he could show how he decided to use an expression rather than an equation (lines 14–15), and he responded contingently, elaborating his decision (lines 17–24). Ms. Castro's response was contingently fitted to his explanation, first summarizing what he did and then asking a further question to prompt an explanation of his next step after the parenthesis was inserted, and so on. Similarly, Ms. Lara and Angie engaged in a series of contingent questions and responses, mostly propelled by Angie's concerns. For example, in lines 19–25 Ms. Lara reads Angie's text and says "“People may argue that,” what is that?” and Angie responded that 'that' was a counter-argument. In the next several lines, both teacher and student build on these observations to create a joint sense of the coherence of Angie's text. In the two teacher–student interactions the contingency of questioning and response resulted in a coherent and continuing conversational sequence.

In a very real sense, these contingent exchanges represent the scaffolding that enables Rico's and Angie's thinking to be externalized and evidenced (Tharp and Gallimore 1988). Indeed, in these exchanges the scaffolding is constructed rung by rung in and through each successive turn at talk and the exchanges that these turns embody. Scaffolding includes enlisting the student's interest in, and adherence to, the requirement of the task, accentuating certain features of the task that are relevant, keeping the student 'in the field' to pursue a particular objective (Wood et al. 1976), and the exchanges described here represent this concerted sharing of a motivational dynamic. In Rico's case, Ms. Castro maintained their focus on the objective of revealing his mathematical thinking related to his problem-solving strategy and provided scaffolding that continuously engaged this thinking throughout their interaction. Similarly, Ms. Lara and Angie were both in pursuit of a resolution to Angie's concern and, through their respective scaffolding, came to point where Angie could take ownership of her next steps in writing through joint activity with her teacher.

As seen above, the interactional sequences that are characteristic of AfL ideally will embody co-regulation between teacher and student in a process in which neither participant is dominant and in which the management of regulation can shift between them. While a primary goal of AfL is to determine the status of learning so as to advance it, co-regulation shapes how this status can be revealed through a process of joint regulatory ownership. As a result, student agency in learning and assessment, a core objective of AfL, is motivated, supported, and upheld by both participants in the interaction.

In the next section, I consider two main challenges to the implementation of interaction in AfL and the implications for co-regulated learning.

## 19.4 Challenges for Implementation

### 19.4.1 *The Social Ecology of the Classroom*

The social ecology of the classroom can either inhibit or enable co-regulated learning. Erickson (2007) observes that in any classroom, ‘proximal formative assessment [assessment for learning] is continually being done from beneath as well as above’ (p. 193). And he continues: ‘As student judgments of and reactions to the teacher influence the conduct of instruction, they are a constitutive feature of the fundamental social ecology of classroom teaching and learning.’ These observations are relevant to both of the teacher–student interactions reported here. Throughout their interactions, from the opening through the pursuit of eliciting student ideas, each teacher upheld the agency and dignity of their students. The interactions took place in a context that was collaborative and non-threatening in nature, and that permitted the students to participate in the AfL process as an equal stakeholder with the teacher. Thus, the assessment culture of the classroom is a contributory factor to the social ecology that supports co-regulated learning.

While many teachers think of themselves as practical and removed or uninterested in the theoretical world, in reality they all have theories that consciously or unconsciously guide their teaching (Heritage et al. 2015). The social ecology of the two teachers’ classrooms is clearly grounded in a socio-cultural perspective (Vygotsky 1986). Students engage with each other and the teacher in a classroom community characterized by participant-oriented learning practices. In this context, student–student and teacher–student interactions are a prime source of learning and of AfL and, as such, they provide participants with opportunities for co-regulation, which are modeled and supported by the teachers. Hadwin et al. (2005) describe self-regulatory ownership and its transition from teacher to student during naturalistic instructional situations. As students are enculturated in the task and task context, teacher regulation should give way to a shared responsibility for regulating learning. This is what is happening in the two classrooms the interactions are drawn from and is enabled by the social ecology of the classroom.

The theoretical perspective inherent in Ms. Lara’s and Ms. Castro’s classrooms contrasts with a stance that is prevalent in many American classrooms: teach, test, and remediate. Teachers teach the content and assessment is carried out at the end of a sequence of instruction to determine students’ level of achievement in ‘in order to fix their failings’ and target the next objective (Klenowski 2009, p. 263). When teachers adopt this approach, which arises from behaviorist theories of learning

where regulation is conducted through reinforcement, students are treated as passive recipients of content, rather than as active participants in a learning or assessment process. Indeed, the term ‘delivering instruction’ is routinely used to describe what happens in classrooms. Contemporary theories of learning that emphasize the active construction of knowledge with others are not reflected in these classrooms, neither in the participant and relationship structures established by the teacher that enable co-regulation in the social ecology of the classrooms.

### ***19.4.2 Teacher Questioning Practices***

As seen above, the nature of the teacher–student interactions in the examples analyzed permitted evidence of learning to be revealed in a context of co-regulation. The questioning patterns adopted by the teachers stand in contrast to the initiation, response, evaluation (IRE) pattern of questioning (Mehan 1979), which still characterizes most questioning practice in American schools (Erickson 2007). In IRE sequences, teachers’ questions are the familiar ‘known answer’ or ‘exam’ questions whose objective is to discover what the student knows (correctly), rather than to engage in a process of understanding students’ thinking, a necessary outcome of AfL questioning. The evaluative nature of teachers’ responses in IRE sequences (right or wrong) instantiate the social relations of the classroom, signaling the epistemic supremacy of the teacher (Heritage and Heritage 2013).

In the two cases examined in this chapter, neither the teachers’ questions—nor in Angie’s case, her own questions—foreclosed the response of the other. Instead, the open-ended nature of the questions, clearly focused on the learning tasks, invited an elaborated response, providing evidence of the students’ thinking. The teachers’ questions were contingent upon the students’ responses. They did not evaluate the students’ responses as in the IRE sequence but rather permitted their students’ thinking to unfold during the sequence. The social relations evidenced in the interactions reflected a partnership between teacher and student, where student agency in the learning and assessment process was both acknowledged and supported by the teacher in the context of the classrooms’ assessment culture. Importantly, the questioning patterns are part of established routines that are recognizable to students, and in which the students and teachers are clear about their respective roles (Allal 2011).

The nature of the questioning practices and the inclusion of student agency that they represent make co-regulation possible. The management of regulation can move back and forth between teacher and student because of the invitations each participant in the interaction provides to the other to engage in understanding thinking in the context of a specific task.

## 19.5 Meeting the Challenges

Recent research has underscored the nature of the challenges for the practice of co-regulation in the context of AfL described above. The Measures of Effective Teaching (MET) project used multiple observation protocols to score the classroom practice of almost 3000 teachers (Kane and Staiger 2012). Findings indicate that teachers demonstrated weak practice in dimensions related to pedagogical and assessment strategies (Kane and Staiger 2012). For example, results from the Classroom Assessment Scoring System (CLASS) observation protocol (Pianta et al. 2008) indicated that less than ten percent of scores were 6 or 7 (on a 7-point scale) for analysis and problem solving, regard for student perspectives, quality of feedback, and instructional dialogue. Clearly this study reflects the need for improved training for teachers in these dimensions of practice which strongly bear on possibilities for co-regulation.

At the preservice level, a focus on implementing pedagogical practices that reflect contemporary theories of learning is necessary. This focus should not be limited to understanding theory but needs to specifically address how these theories are instantiated in daily classroom practices (Otero 2006). Preservice teachers need opportunities to engage in careful analyses of how the social ecology of the classroom is formed to support the students as agents in their own learning. Also needed is a focus on dialogue and as well as the nature of questioning and responses that permit co-regulation in learning. Some promising practices are beginning to emerge in this area (see for example, Duckor 2014), but much more remains to be done to ensure that beginning teachers are equipped to ask questions, listen carefully to their students, and respond contingently as their thinking is revealed.

Beyond preservice education, continuing professional learning could fruitfully engage teachers in deep analysis of practice that centers on questioning sequences that enable teachers to obtain evidence of student learning within a context of co-regulation. In U.S. schools, ‘close-to-practice learning,’ where teachers analyze video of their and others’ practice, is becoming more common (see for example, Ermeling and Gallimore 2014). School leaders and teachers could prioritize this kind of analysis in professional learning in order to consider the theoretical perspectives reflected in the pedagogical and learning practices. Video analysis can also help in understanding the social ecology of the classroom, the questioning and response patterns that enable learning, that constitute AfL and permit co-regulated learning. A focus on the environment and routines will necessarily involve developing teachers’ general classroom management skills, including how to organize the classroom so that students can work independently while the teacher engages in one-to-one or small-group interactions.

In terms of policy, the practices exemplified by Ms. Castro and Ms. Lara could be more squarely placed as core to effective teaching in the current push in the U.S. for stronger teacher evaluation frameworks. More fine-grained attention to classroom interaction would accomplish two goals: (1) highlight the value of these interactions as a component of teaching and learning; and (2) provide opportunities



for feedback from evaluators that assist teachers to strengthen their practice making co-regulation a possibility in their classrooms.

The above recommendations would not be financially costly to implement. Granted, there would likely be a cost to building the capacity of teacher leaders or professional development providers so they could offer the necessary support for teachers. However, compared with the extensive sums of money many governments are willing to spend on large-scale testing, this cost may be relatively minor. Moreover, such expenditure would create a human infrastructure with the ability to stimulate teachers' thinking and practice focused on obtaining the evidence they need to keep learning moving forward and assisting students in internalizing self-regulatory progresses as a result of co-regulated interaction with their teachers.

## References

- Allal, L. (2010). Assessment and the regulation of learning. In P. Peterson, E. Baker, & B. McGraw (Eds.), *International encyclopedia of education* (Vol. 3, pp. 348–352). Oxford: Elsevier.
- Allal, L. (2011). Pedagogy, didactics and the co-regulation of learning: A perspective from the French-language world of educational research. *Research Papers in Education*, 26, 329–336.
- Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, 50, 248–287.
- Beach, W. A. (1993). Transitional regularities for casual 'okay' usages. *Journal of Pragmatics*, 19, 325–352.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. New York: Open University Press.
- Black, P. J., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice*, 5, 7–73.
- Black, P., & Wiliam, D. (2005). Changing teaching through formative assessment: Research and practice. In *Formative assessment: Improving learning in secondary classrooms* (pp. 223–240). Centre for Educational Research and Innovation (CERI). Paris: OECD.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability*, 21, 5–31.
- Calkins, L. M. (1994). *The art of teaching writing*. Portsmouth, NH: Heinemann.
- Duckor, B. (2014). Formative assessment in seven good moves. *Educational Leadership*, 71(6), 28–32.
- Duranti, A. (2010). Husserl, intersubjectivity and anthropology. *Anthropological Theory*, 10(1–2), 16–36.
- Erickson, F. (2007). Some thoughts on 'proximal' formative assessment of student learning. *Yearbook of the National Society for the Study of Education*, 106, 186–216.
- Ermeiling, B. A., & Gallimore, R. (2014/2015). Close-to-practice learning. *Educational Leadership*, 72(4). Retrieved from [http://www.ascd.org/publications/educational\\_leadership/dec14/vol72/num04/Close-to-Practice\\_Learning.aspx](http://www.ascd.org/publications/educational_leadership/dec14/vol72/num04/Close-to-Practice_Learning.aspx)
- Greeno, J. G. (2006). Learning in activity. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 79–96). New York: Cambridge University Press.
- Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 15–46). New York: Macmillan.

- Hadwin, A. F., Järvelä, S., & Miller, M. (2011). Self-regulated, co-regulated, and socially shared regulation of learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *The handbook of self-regulation of learning and performance* (pp. 65–84). New York: Routledge.
- Hadwin, A., Wozney, L., & Pontin, O. (2005). Scaffolding the appropriation of self-regulatory activity: A socio-cultural analysis of changes in teacher-student discourse about a graduate research portfolio. *Instructional Science*, 33(5–6), 413–450.
- Heritage, J. (1984). *Garfinkel and ethnomethodology*. Cambridge: Polity Press.
- Heritage, J. (2010). Questioning in medicine. In A. F. Freed & S. Ehrlich (Eds.), *Why do you ask? The function of questions in institutional discourse* (pp. 42–68). New York: Oxford University Press.
- Heritage, J. (2013a). Epistemics in conversation. In J. Sidnell & T. Stivers (Eds.), *Handbook of conversation analysis* (pp. 370–394). Boston, MA: Wylie-Blackwell.
- Heritage, M. (2013b). *Formative assessment in practice: A process of inquiry and action*. Cambridge, MA: Harvard Education Press.
- Heritage, M., & Heritage, J. (2013). Teacher questioning: The epicenter of instruction and assessment. *Applied Measurement in Education*, 26, 176–190.
- Heritage, M., Walqui, A., & Linqunti, R. (2015). *English language learners and the new standards*. Cambridge, MA: Harvard Education Press.
- Järvelä, S., & Järvenoja, H. (2011). Socially constructed self-regulated learning and motivation regulation in collaborative learning groups. *Teachers College Record*, 113(2), 350–374.
- Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teachers: Combining high-quality observations with student surveys and achievement gains*. Seattle, WA: Bill and Melinda Gates Foundation.
- Klenowski, V. (2009). Assessment for learning revisited: An Asia-Pacific perspective. *Assessment in Education: Principles, Policy and Practice*, 16(3), 263–268.
- Mehan, H. (1979). *Learning lessons*. Cambridge, MA: Harvard University Press.
- Otero, V. (2006). Moving beyond ‘get it or don’t’ conception of formative assessment. *Journal of Teacher Education*, 57, 247–255.
- Pianta, R. C., Paro, K. M. L., & Hamre, B. K. (2008). *Classroom assessment scoring system (CLASS) Manual, K–3*. Baltimore, MD: Paul H. Brookes Publishing.
- Raymond, G. (2003). Grammar and social organization: Yes/No interrogatives and the structure of responding. *American Sociological Review*, 68, 939–967.
- Ruiz-Primo, M. A., & Furtak, E. M. (2006). Informal formative assessment and scientific inquiry: Exploring teachers’ practices and student learning. *Educational Assessment*, 11, 205–235.
- Sacks, H. (1987). On the preferences for agreement and contiguity in sequences in conversation. In G. Button & J. R. E. Lee (Eds.), *Talk and social organization* (pp. 54–69). Clevedon, UK: Multilingual Matters.
- Sadler, D. R. (1989). Formative assessment and the design of instructional strategies. *Instructional Science*, 18, 119–144.
- Schegloff, E. A. (1992). Repair after next turn: The last structurally provided for place for the defense of intersubjectivity in conversation. *American Journal of Sociology*, 95(5), 1295–1345.
- Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis* (Vol. 1). Cambridge, UK: Cambridge University Press.
- Shavelson, R. J., Young, D., Ayala, C., Brandon, P., Furtak, E., Ruiz-Primo, M. A., et al. (2008). On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied Measurement in Education*, 21(4), 295–314.
- Smith, J. (2003). Reconsidering reliability in classroom assessment and grading. *Educational Measurement: Issues and Practice*, 22(4), 26–33.
- Swaffield, S. (2011). Getting to the heart of authentic assessment for learning. *Assessment in Education: Principles, Policy and Practice*, 18(4), 433–449.
- Tharp, R., & Gallimore, R. (1988). *Rousing minds to life*. New York, NY: Cambridge University Press.

- Torrance, H., & Pryor, J. (1998). *Investigating formative assessment*. Buckingham, UK: Open University Press.
- Vygotsky, L. (1986). *Thought and language*. Cambridge, MA: MIT Press.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Psychology and Psychiatry*, 17(2), 89–100.

## Chapter 20

# Supporting Students' Learning: From Teacher Regulation to Co-Regulation

Louise Bourgeois

**Abstract** Even though there is widespread acceptance that student interaction with peers could make a strong contribution to learning, there is evidence that teachers have trouble adopting formative assessment strategies that actively involve students. They often see activities like peer feedback as less efficient than direct instruction and are frustrated by the results of such co-regulated activities. They are, however, also disappointed with how students generally respond to teacher feedback which may be an indicator of problems either with the quality of the feedback or with the students' capacity to regulate their own learning. The qualitative multiple case study research presented in this chapter was designed to understand what formative assessment decisions teachers make to support learning and to help them move progressively from teacher-focused to student-centered action. The research is focused on language arts, more specifically on writing, and it is divided into three phases. At each phase, teachers select and share texts produced by their students with a colleague and make individual and joint assessment decisions to support student learning. The first phase of the research is designed to understand the assessment decisions teachers are initially making to support learning. Each of the following phases begins with an intervention in order to determine its impact on teachers' assessment decisions. Findings reveal that changes in formative assessment practice do occur as a result of discussions with the colleague and the interventions, however, they occur unequally among teachers and are influenced by their individual representations of what formative assessment involves in practice.

---

L. Bourgeois (✉)  
University of Ottawa, Ottawa, Canada  
e-mail: lbour028@uottawa.ca

L. Bourgeois  
Laurentian University, Sudbury, Canada

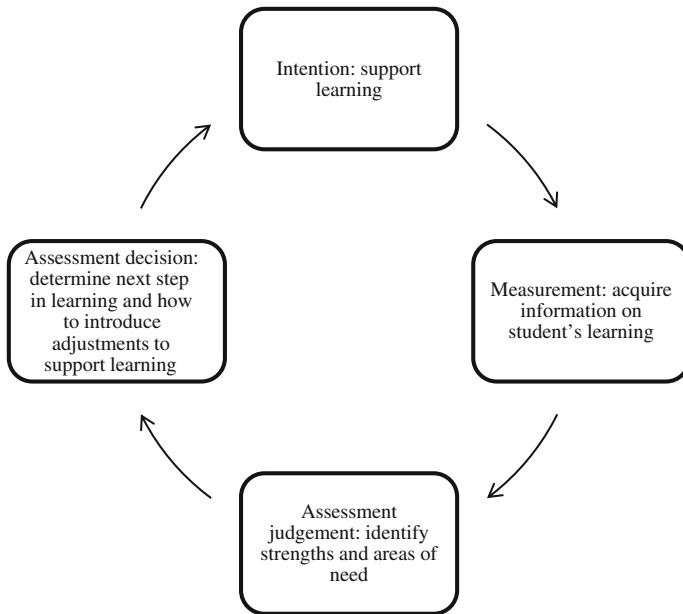
## 20.1 Introduction

Although they share similar definitions, the terms ‘formative assessment’ in English and ‘*évaluation formative*’ in French have fundamental differences in perspective. In the English-language literature, formative assessment (FA) is closely associated with feedback and remediation while in the French-language literature, it is much more closely linked to the notion of *régulation* (Allal and Mottier Lopez 2005; Laveault 2006; Wiliam 2011), which includes a diversity of adjustments to the processes of teaching and learning. In French-language research, the concepts ‘regulation of teaching’ and ‘regulation of learning’ are frequently espoused. The notion of regulation as being central to FA appears, nevertheless, implicitly in some English-language research. For instance, Heritage et al. (2009) show that while teachers may be good judges of the strengths and weaknesses of students’ work, they less often concur as to the next steps in instruction to help support student learning. This study shows that the real challenge for teachers in FA does not appear to be teachers’ capacity to make reliable judgments on the quality of student work but rather their capacity to use the information gathered on student learning to make pertinent assessment decisions about the adjustments to be undertaken. The intervention component which assures effective regulation appears to be the point at issue for teachers in FA. As Heritage et al. (2009) mention, ‘This situation inevitably diminishes the potentially powerful impact of formative assessment on student learning’ (p. 31).

The research presented in this chapter focuses on teachers’ assessment decisions in a formative context, more specifically, it focuses on the types of regulation they introduce to improve learning, and examines what can be done to support their capacity to use information on learning to better respond to students’ learning needs in the area of writing. In the formative assessment process (Fig. 20.1), assessment decisions first involve acquiring information on students’ learning and then determining strengths and areas of need. The assessment decision itself entails identifying the next step in learning for a student (e.g., use adjectives to make the text more vivid) and then determining the intervention that will help move the student forward (e.g., have the student work with a peer to identify nouns in the text and find appropriate adjectives).

In recent years, it has become common to describe the use of assessment to improve student learning as ‘assessment for learning.’ In this chapter, the terms formative assessment (FA) and assessment for learning (AfL) are used interchangeably and are aligned with the following definition:

The process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there. (Assessment Reform Group 2002, p. 2)



**Fig. 20.1** Formative assessment process

## 20.2 A Typology of Regulation

The notion of regulation involves four fundamental components (Allal 2010; Laveault 2000, 2006):

- a goal or a standard to help determine the learning target;
- a form of control to help determine how well the learning is progressing towards the goal;
- a judgment on strengths and weaknesses to focus on what needs to be improved;
- an adjustment or other action to help guide the next steps in learning.

Allal (1988) proposes three types of regulation: proactive, interactive, and retroactive regulations. In fact, these modalities refer to when adjustments are made. Proactive regulations are made ‘upstream’ of the teaching situation. For example, in the planning of a lesson, a teacher may use experiential knowledge of students’ learning needs to plan the best course of action required to effectively support learning. Interactive regulations occur during the teaching and learning activity. For example, when a teacher uses students’ responses to questions and prompts to try to establish whether students have understood what they are meant to be learning and to adjust the course of a lesson accordingly. Retroactive regulations entail a remediation that occurs at the end of the lesson when, for example, a teacher collects information which provides evidence of learning to help plan corrective instruction. Effective teaching generally includes all three modalities of regulation

(Allal and Mottier Lopez 2005) where retroactive regulations help remediate difficulties that persist after proactive and interactive regulations have run their course (Allal 1988).

Not only can these modalities of regulation be used by the teacher to adjust teaching but they can also be used by students to adjust their own learning. In fact, Scallon (2000) suggests that FA hinges upon two possible options: the regulation of teaching, where the teacher makes adjustments to the teaching situation (external regulation), and the regulation of learning, where students make adjustments to their learning (internal regulation). It is worth noting that the teacher cannot regulate student learning directly but he or she can support the student in regulating his or her own learning through co-regulation (Allal 2010). In that respect, students can make proactive regulations, for example when they refer to assessment criteria to determine how to carry out a task, interactive regulations when they use feedback to adjust their work as they go along, or retroactive regulations when they make changes to a draft through self-assessment using the assessment criteria.

### ***20.2.1 Regulation Failures***

Regulations whether of teaching or learning can also miss the mark. Baumeister et al. (1994) identify three types of regulation failures:

- overregulation when too many adjustments occur at once or from too many sources;
- underregulation when too few or no adjustments occur when they should;
- misregulation when adjustments are misdirected.

Along with the difficulty of ensuring that regulations occur at the right time, in the correct dosage, and on the appropriate objects, Perrenoud (1998) adds that the regulation of teaching in itself is a difficult process for the teacher to implement. It is highly demanding because assessment decisions are made in a context of uncertainty where information on learning is either incomplete or missing. He explains:

even a teacher possessing all the theoretical tools would have considerable difficulty using them in an optimal way in a given situation, because of other priorities and uncertainties which cannot be dealt with in a given time. (p. 87)

Boekaerts (1997) also suggests that the regulation of teaching alone can be ineffective not only because the learning process is difficult to control externally but because it can dampen students' engagement and motivation to participate in the instructional process:

External regulation is a form of support that leaves the learner little autonomy and hardly any responsibility for the learning process. The scaffolding metaphor captures the idea of an adaptable and temporary support system that helps an individual during the initial period of gaining expertise. (p. 171)

In this passage, Boekaerts highlights the concept of scaffolding where the teacher enables students to do with help that which they would not be able to do alone. However, the implied process of ‘unscaffolding’—where teacher support is progressively removed as students gain knowledge and skills—is lesser understood and much more difficult for teachers to implement (Meirieu 2006). The challenge for teachers may be knowing how to go about teaching students to self-regulate and then progressively tilting the balance of control to the student. In fact, this entails the active involvement of students in FA through self-assessment and self-regulation.

### 20.2.2 Supporting Self-Regulation

To a great extent, self-regulation involves much the same processes as teacher regulation. Students must acquire similar skills to those of their teacher and be able to set goals, monitor progress towards the goal, make judgments on the quality of their work, and adjust learning (Sadler 1989). Self-assessment is central to self-regulation and the goal in this context is to support students’ awareness of the quality of their work and their capacity to monitor their progress. In the end, supporting students’ self-regulation skills is not so much about supporting learning directly but rather about supporting the student as a learner (Laveault 2014).

Stiggins et al. (2004) refer to FA as addressing three main questions: Where am I going? Where am I now? How am I going to get there? Based on these questions and on the four fundamental components of regulation presented above, a fourth question can be added to show how teacher and student can share in the responsibility of assessment decisions (see Table 20.1).

For example, the teacher alone could be responsible for setting the learning goal while the student could be responsible for monitoring his or her progress towards that goal. Both teacher and student could share in the responsibility of making judgments on the quality of the work (co-evaluation), as well as determining next steps in learning and improvement strategies (Laveault 2007). As students’ self-assessment and self-regulation skills progress, the goal is to gradually give them more control over the regulation of learning.

**Table 20.1** Who makes assessment decisions?

Teacher regulation	Assessment decision	Student regulation
Where is the student going?	Setting goals	Where am I going?
How is the student doing?	Monitoring progress towards the goal	How am I doing?
Where is the student now?	Making judgments on the quality of the work	Where am I now?
How is the student going to get there?	Making adjustments to support the next steps in learning	How am I going to get there?



Co-regulation can be an effective way to progressively move students from teacher-focused to student-centered action and self-regulation of learning. Hadwin and Oshige (2011) provide a definition of co-regulation that explains the iterative process between external and internal regulations that occurs between student and teacher, or among students, to gradually support self-regulation skills:

Coregulation refers to a transitional process in a learner's acquisition of self-regulated learning (SRL), within which learners and others share a common problem-solving plane, and SRL is gradually appropriated by the individual learner through interactions. (p. 247)

Interaction with peers can be an effective stepping-stone to help students acquire the self-assessment skills needed to make better assessment decisions (e.g., Allal 2011; Hadwin and Oshige 2011). For example, reciprocal peer assessment where each student takes the role of both assessor and assessee can help students develop internal standards for quality work and support their capacity to make better judgments of their own work.

The notion of regulation is like a two-arm balancing scale where the regulation of teaching falls at one end, the regulation of learning at the other, and co-regulation works at tipping the scale towards self-regulation of learning. For example, at the beginning of a learning unit, the scale is loaded on the side of teacher regulation as students are just starting to acquire knowledge and skills. However, as student learning and assessment skills progress, for example, through co-regulated activities (e.g., co-construction of criteria, peer feedback), the teacher can progressively scale down on the control of one or more of the assessment decisions and students can gradually exercise more control on regulating their learning. Ultimately, the goal is to give increasingly more control to the student of most, if not all of the assessment decisions (Laveault 2007).

### 20.3 A Study of Teachers' Decisions About Regulation

This research was designed as a multiple case study that employs qualitative methods. It was intended to better understand what formative assessment decisions teachers make to support learning and to help them move progressively from teacher-focused to more student-centered action. In this chapter, two cases are documented and differences within and between both cases are explored. Each participant is given a pseudonym to ensure anonymity.

Six grade 7 and 8 teachers from French-language schools in Ontario participated in the research. In Ontario, the 7th and 8th grades generally represent the first years of secondary school; however, students at these grade levels (students aged 11–14) are not yet academically streamed. Teachers were selected by convenience sampling using selection criteria; they had to be qualified teachers with a minimum of five years teaching experience and responsible for teaching language arts (*français*) in grade 7 and/or 8. The two case studies presented in this chapter were selected because they represented typical examples of the teachers studied.

The research focuses on two central questions:

- What type of regulation does the teacher determine to be the best course of action to support the student's next step in learning in writing?
- What changes in type of regulation can be observed across the three phases of the research for a given teacher?

### **20.3.1 Method**

The research was divided into three phases (see Table 20.2) that spanned the length of a school year (2013–14). The data collection tools included:

- a questionnaire;
- a retrospective semi-directed individual interview;
- direct observation of discussions between the teacher and the colleague;
- a semi-directed interview involving the teacher and the colleague.

At each phase, teachers selected two texts produced by students in their class and shared them with a colleague. Prior to a retrospective semi-directed interview, each teacher responded individually in writing to a questionnaire that consisted of five open-ended questions concerning the student's text:

1. What are the strengths?
2. What are the areas of need?
3. What is the next step in learning for the student?
4. What is the best course of action to support the student's next step in learning?
5. What written feedback would you give the student?

The questionnaire was largely based on questions Heritage et al. (2009) asked teachers in their generalizability study on the measures of teacher knowledge for teaching mathematics.

The first phase of the research was designed to better understand what the teacher was initially doing to support learning. Each of the following two phases included an intervention in order to determine its impact on teachers' assessment decisions. Prior to the second phase, teachers read the provincial policy on assessment, more specifically the chapter on FA, and set professional learning objectives for themselves. At the beginning of the third phase, teachers worked with students on a series of activities intended to develop students' self-assessment and peer-feedback skills. Following the intervention for phases 2 and 3, the teacher followed the same process of assigning a type of text for the students to write, selecting and sharing texts with the colleague, responding to the questionnaire, and participating in a retrospective interview, a discussion with the colleague, as well as a semi-directed interview with the colleague.

**Table 20.2** Three phases of the research on teachers' decisions about regulation

	Phase 1 Oct. to Dec. 2013	Phase 2 Jan. to March 2014	Phase 3 Apr. to June 2014
Intervention	• No intervention	The teacher • read the provincial assessment policy; • set professional learning objectives.	The teacher • worked with students on a series of activities intended to develop their assessment skills.
Teacher actions	The teacher • taught knowledge and skills related to a type of text (e.g., detective story); • selected texts from two students; • shared texts with the colleague; • answered the questionnaire for each text; • participated in a retrospective individual semi-directed interview; • discussed and adjusted assessment decisions with the colleague; • participated in a semi-directed interview with the colleague.		

### **20.3.2 Provincial Policy on Assessment: Phase 2 Intervention**

In the area of FA, the Ontario policy on assessment highlights the importance of engaging students in the assessment process (Ministry of Education 2010). In fact, the policy mandates the use of five strategies to support FA which include:

- Developing and sharing learning targets and success criteria with students (e.g., co-constructing criteria with students);
- Elaborating effective learning tasks that generate targeted information about student learning;
- Providing descriptive feedback that moves students forward;
- Providing targeted instruction to help students become learning resources for one another (e.g., peer-feedback skills);
- Providing targeted instruction to help students take ownership of their learning (e.g., self-assessment skills).

Prior to phase 2 of the research, teachers read the chapter on FA of the policy and set professional learning objectives for themselves. It was expected that teachers would focus on one or more of the mandated strategies that they felt they needed to improve upon to support more effective teaching and learning. It was also expected that this focus would in turn foster some fundamental changes in teachers' FA practice and assessment decisions (e.g., identifying clear assessment criteria, using assessment information on student learning to plan instruction). At the time of this research, the implementation of the policy had run three years of its course.

### 20.3.3 Activities to Support Students' Assessment Skills: Phase 3 Intervention

The activities that were developed to support students' assessment skills consisted of a student booklet and a teacher's guide (Laveault et al. 2013). At the beginning of phase 3, teachers used the activities in the booklet with their students. Although the explicit intention of the booklets was to support students' self-assessment and peer-feedback skills, the implicit intention was to support teachers':

- capacity to improve their feedback skills;
- capacity to use peer and self-assessment more effectively with students;
- understanding of the intrinsic relationship among the five strategies mentioned in the provincial assessment policy (see Sect. 20.3.2);

Figure 20.2 gives an overview of the activities contained in the booklet. It was expected that teachers' use of the booklet with their students would foster changes in their beliefs about involving students as learning resources for one another and in the focus of their assessment decisions (e.g., moving from teacher-focused to student-centered action).

The activities in the booklet focused on a descriptive text: the summary. The first activity involved students individually writing a summary of an article. Through guided practice, students were then led through the co-construction of assessment criteria for the summary and after that, in small groups, they were asked to

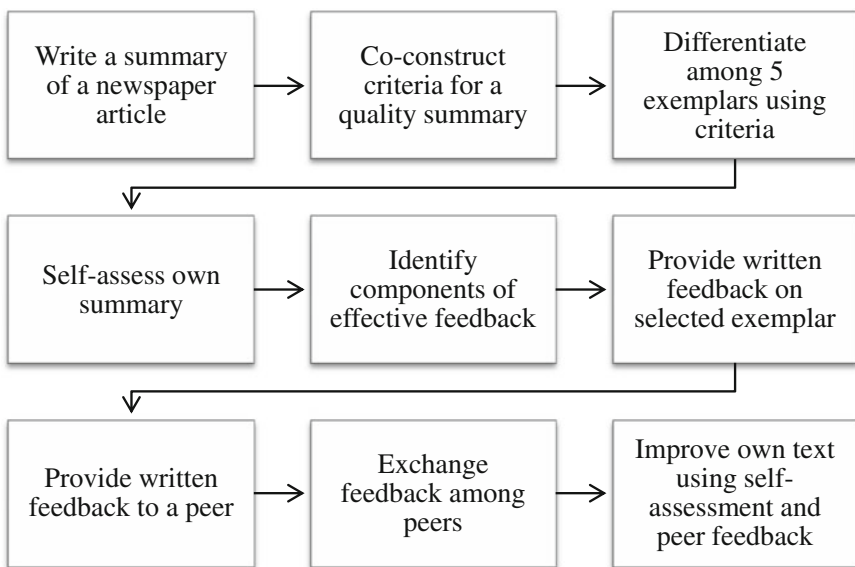


Fig. 20.2 Progression of activities in the student booklet

differentiate among five exemplars<sup>1</sup> of summaries of the same article using the criteria they developed. Students had to assign a level of achievement to each exemplar—from level 1 that represents a weak text to level 4 that represents an excellent text—and explain their reasoning. The goal was to help students develop internal standards for quality work and support their capacity to make better judgments about their own work. Next, using the assessment criteria, students assessed their own summary. Students were later asked to read two examples of feedback for the level 1 exemplar and to select the feedback that was most effective. Subsequently, students determined the essential components of effective feedback (e.g., identify strengths and weaknesses, suggest a strategy to improve upon a weakness) and practiced their newly honed skill on the level 2 exemplar. In dyads, students then exchanged the summaries they had written at the beginning and wrote feedback for each other. Finally, students used the feedback provided by their partner to improve their summary.

### **20.3.4 Data Analysis**

For each of the three phases of the research, the data collected for a teacher consisted of students' texts, responses to the questionnaire concerning each text and the written transcription of the retrospective semi-directed interview, the discussion between the teacher and the colleague, as well as the semi-directed dyad interview. A qualitative content analysis was conducted on the written transcriptions for each teacher by extracting the segments that were related to FA (e.g., representation of FA, classroom strategies). When teachers referred to the student's text during the interviews or the discussion (e.g., to further explain the regulation mentioned in the questionnaire), those segments were extracted and added to the data in the questionnaire to be analyzed as such. Content analysis was performed on the data collected from the questionnaire. The main purpose was to determine what changes, if any, occurred in the teacher's assessment decisions following the discussion with the colleague and each intervention. To ensure the reliability of the content analysis, intercoder agreement was used and discrepancies were discussed and resolved.

## **20.4 Case Studies**

The case studies presented here concern two 7th grade teachers from French-language schools in Ontario. Both are women with approximately 10 years of teaching experience who have identical teaching assignments (three language arts classes) in different schools and approximately the same number of students

---

<sup>1</sup>Exemplars are examples of student texts illustrative of different levels of performance in writing.

(90 students). The teachers had both been involved in a wide range of professional development activities in language arts (*français*) and assessment throughout the years. Both teachers had discussions with a colleague in their respective schools. Although the cases show contrasts with respect to individual representations of what FA involves in practice, it could be argued that both cases are typical examples of the implementation of FA in the classroom. Each case study describes the journey of the teacher with the students of her class through the three phases of the research.

### 20.4.1 Case Study 1: The Journey of Mireille and Her Class

At the outset of phase 1, Mireille's description of the FA practices she uses in her classroom mostly involves strategies geared towards teacher regulation.

This year, I'm using exit tickets<sup>2</sup> in grammar and that's working out really well. So, I'll teach a grammatical concept and then students do the exit ticket and give it back to me before they leave so I can plan for the next day. It tells me who got it and who needs more examples. It's very quick, one or two questions about what I just taught and I know who I have to work with the next day. I'll give it back to them with a comment like 'You need to review personal pronouns.' Then, we'll correct it together in class and I'll give more examples to help the students who didn't get it right. Sometimes, I'll take those students aside and give them extra help and reteach the concept.

I also ask a lot of questions when I'm teaching and I use the popsicle sticks<sup>3</sup> with the names of the students on them. It's a great way to keep them listening because even if they don't raise their hand, they never know when I'll call on them. It keeps their attention on the lesson.

I'll also give them feedback. While they are working on the draft, I'll meet with them individually and we'll work on the text together. Then, I'll give them one thing they need to improve and they'll go back and work on that. Sometimes, I'll circle mistakes in their text or I'll underline a sentence that is not clear and they have to figure out how to correct it.

I give them a checklist of things to check like spelling, punctuation, sentence structure. That I find doesn't work very well. They tend to just put check marks all the way down the column without checking their work. I give them a lot of handouts like that but they don't use them. I don't know why, maybe it's too much. But even written feedback, they don't use unless I sit down and explain it to them one item at a time. I don't know if it is because they don't care about improving their work or they don't understand the feedback, maybe it's a little of both.... I used to do peer review in writing but now I don't use it very often

---

<sup>2</sup>An exit ticket is a prompt or a question that is given to students at the end of a class to provide feedback to the teacher about what students have learned and to help plan the next lesson. It usually requires only a brief amount of time for students to complete and it is handed back to the teacher as students leave the class or transition to another subject.

<sup>3</sup>Students' names are written on the wooden popsicle sticks that are placed in a cup. During classroom questioning the teacher draws a stick at random to determine who will answer. This technique can ensure that all students have an opportunity to answer during class and that consistent hand-raisers are not dominating classroom interaction.

because it's not really useful. You look at the revisions they make after they worked together and it's not worth the time it took.... I'm not sure about things like co-constructing the criteria with students. I struggle with developing good criteria myself so how can I expect them to be able to do it.

Mireille's focus is mostly on strategies related to the teacher regulations (e.g., exit tickets, reteaching a concept). It is also possible to detect some elements of overregulation in her practice, related to the number of handouts she gives students with the intention of supporting learning. As she mentions, this practice may have the adverse effect of overwhelming students rather than helping them. Mireille also gives feedback to support learning but she is disappointed with how students respond to it, which may be an indicator of problems with the quality of her feedback or with the students' capacity to regulate their own learning. Her experience with co-regulated activities such as peer review has not been successful and she shies away from collaborative interaction among peers to construct meaning and understanding, preferring instead to use more teacher-directed strategies such as direct instruction. She does however expect students to self-regulate (e.g., use a checklist to revise) without much guidance or support. Although she is aware of the push toward student involvement in the assessment process to support learning, she is not convinced that such strategies (e.g., co-constructing assessment criteria) are effective or even viable.

Through her participation in the research, it is possible to track specific changes in her assessment practice where she progressively attempts to move from teacher-focused to more student-centered regulations.

- By the end of phase 2, Mireille was developing assessment criteria in a language that was accessible to students, sharing the criteria with students and using her own exemplars to support their understanding of the criteria. The criteria she had initially developed in phase 1 were not written in a language that was accessible to students (e.g., 'The student demonstrates understanding of the characteristics and the structure of the type of text.') and were never explicitly shared with them. During phase 3, she made a first attempt at co-constructing criteria with students after having done a similar activity with the student booklets.
- At the beginning of phase 1, Mireille was meeting regularly with individual students at the drafting stage of the writing process and identifying spelling and grammar errors for them to correct. In phase 3, she tweaked this approach, first asking students to identify spelling and grammar errors they thought they might have made, as a way of obtaining more information about student understanding. She was then able to use the information to better support learning.
- In phase 1, Mireille's assessment decisions as to the best course of action to support students' next step in learning were either geared toward teacher regulations (e.g., remedial instruction for the student) or instructing students to review a handout sheet to improve an element of writing (e.g., 'Review the handout on discourse markers to support text fluidity.'). In phase 2, through discussions with her colleague, she considerably reduced the number of teacher handouts. Together, Mireille and her colleague proposed a student-developed

reference tool they named 'the placemat'. It begins as a single blank page at the beginning of a learning unit that the students gradually 'fill out' as they acquire the knowledge (e.g., characteristics of the type of text) and skills (e.g., how to use organization markers) needed to write a quality piece of writing. Although Mireille gave students full control of this reference tool, she also showed them examples of 'good' placemats. The students' set the placemat on their desk at the beginning of every lesson and referred to it as required. During phase 3, Mireille mentioned that she would also encourage students to determine by themselves when they felt it was appropriate to stop using the placemat.

- During phase 3, Mireille was also using some co-regulated activities with peers to support learning (e.g., asking students, in groups, to improve a text). Although external regulation was still the mainstay of her assessment decisions in phase 3, the balance of power had somewhat shifted in the classroom and she was providing targeted instruction to help students become learning resources for one another (e.g., asking peers to read each other's texts and underline a sentence they found difficult to understand).
- From the outset, Mireille was providing regular oral feedback to students. Although it was not possible to analyze the content of this feedback, content analysis of the written feedback from her answers to the questionnaire from phase 1 showed that she informed students of the strengths and weaknesses of their text; however, proposals of improvement strategies were either absent from the feedback or too vague to effectively support students (e.g., 'Review the example of the introduction given in class to improve yours.'). After using the student booklets in phase 3, improvement strategies were more specific (e.g., 'Make sure you give a few hints at the beginning of your story about who is responsible for the crime.').
- Mireille also used the feedback template from the student booklet to give written feedback to students during phase 3 when she taught the detective story. Students were familiar with the feedback template having used it themselves to give feedback to one another during the activities with the booklets. She noticed that the template appeared to support students' understanding and interpretation of the feedback, and she did not need to meet with as many students individually to explain the feedback.

### ***20.4.2 Case Study 2: The Journey of Céline and Her Class***

At the outset of phase 1, Céline's description of the FA practices she used in her classroom was closely connected to her concern with summative assessment.

For me, everything I do until the summative assessment, the test, is formative assessment. You know, everything I teach, all the activities I plan, that's all formative assessment.... It's an opportunity for students to prepare for the summative assessment. Let's say I'm teaching a type of text, the structure of that text, they'll read the text and I'll ask them questions



about the structure. Those questions I ask while I'm teaching are very similar to the questions that will be on the test. So it prepares them. Then we'll correct their answers. They'll write them on the board and we'll compare them and I'll say, if you answer this way on the reading test, that's a level 3 but if you answer this way, that's a level 4, and then they have to tell me how the two answers are different. Why one is better than the other.

In writing, it's much the same thing. Let's say I'm teaching the opinion text, we'll start by reading some together to identify the characteristics of that type of text and the structure too. Then they'll read some individually to make sure they can do it on their own. We'll correct it in class so they'll know if they understand or not. I let them know that the characteristics and the structure are extremely important because if you don't respect all the characteristics and the structure when you write your text, well then you won't even be able to get a level 3.

I don't really get self-assessment. I don't see the point or how it helps students. It doesn't help me to know that the student thinks his work is a level 3 or a level 2.... In the end, assessment is my responsibility.

Feedback, that's mandated in our school board. It says that we always have to have an achievement chart with every piece of work that counts for the report card and we have to write a personalized comment at the bottom. You know, strengths, weaknesses, and next steps. It takes a lot of time because you have to pinpoint all those things in the student's work, but it's good because it keeps the parents informed.

When asked if she offered feedback to students on formative writing tasks as well, Céline responded:

In a perfect world, it would be great to do that, you know, to give feedback on formative assignments but it's just too time-consuming. In our school board, we have to have at least two summative assessments in each of the three strands in *français* for every report card. That's six summative assessments, twice a year. I have 90 students in *français* so if you do the math, that's close to 1200 summative assessments a year that I have to correct and write a comment. If you consider that each summative assessment takes about 15 minutes to correct, that's like 300 hours of work outside of the school day. I don't have the time to do more than that. What we do though for the students that are at risk, those who perform at level 1, we have a period once a week where we can withdraw the student from another class and give him or her extra help one-on-one or in a small group. So that's what I'll do, I'll go through the test or the assignment with the student and I'll show him or her what he can do to get to a level 2 next time.

Céline's focus on summative assessment made it difficult to determine what type of regulation she used to support learning in writing. In fact, once the students wrote a text, they were immediately moved into a summative process where they received feedback without the opportunity to use it to improve the text. Moreover, Céline does expect students to self-regulate, that is, to use the feedback from one summative assessment to improve on the next one, or to use information from the collective correction of homework to determine what they understand. Such feedback does little to support students' focus on a specific learning goal and by the same token may not lead to improvements in learning.

Through Céline's participation in the research, it is possible to track some changes in her FA practice.

- At the beginning of the research, the assessment criteria available to determine the quality of the text were very narrow in scope focusing mostly on language conventions (e.g., spelling, grammar) and in some cases corresponded to instructions on what to include in the text (e.g., 'Include an introduction and three arguments'). So it was fitting that after reading the provincial assessment policy, Céline chose to work on criteria as her professional learning objective. During phase 2, the criteria were wider in scope (e.g., quality of ideas, organization of ideas), but it was not until phase 3 that assessment criteria were written in a language that was accessible to students (e.g., 'Use a variety of sentences, such as interrogative, exclamatory, imperative and declarative sentences'). During phase 2 and 3, she presented the criteria to students without targeted instruction to support their understanding and without following through at the end with a peer-assessment or self-assessment activity for students. Nevertheless, she felt that some students had benefitted from being informed of the assessment criteria.
- During phase 3, Céline did attempt to implement what she called 'semi-formative assessment' where she still assessed the text summatively by assigning a grade but gave students the choice to improve their text using the feedback with the possibility of getting a better grade. This had limited success as very few students took her up on it. She believes, however, that if she made the activity mandatory, students' texts would improve.
- While the scope of the assessment criteria did widen during phase 2, assessment decisions on the next steps in learning remained focused on language conventions throughout the three phases: 'Explain to the student the importance of correcting spelling and grammar errors in her text'; 'Model how to correct errors using a student's text'; 'Show students how to use resources such as the dictionary and the *Bescherelle* to make corrections.' The course of action to support the student's learning was teacher regulated throughout the three phases of the research. Céline mentioned that she did implement some of the regulations of teaching she had indicated in the questionnaire and was hopeful that they would have an impact on students' revision skills in the future.
- Phase 1 feedback consisted in identifying strengths and areas of need within the students' texts. Some progress could be seen in the feedback from phases 2 and 3; however, proposed improvement strategies were generally vague or difficult for the students to use by themselves (e.g., 'Review the characteristics and the structure of the text we read in class and make sure that you do the same in your text.').
- During phase 1 of the research, discussion with her colleague did generate assessment decisions that improved upon Céline's initial decisions, shifting the focus away from revising grammar and spelling mistakes to, for example, improving text structure. Céline was reticent to accept her colleague's ideas, stating that she 'wasn't there yet.' During phases 2 and 3, Céline did attempt to integrate some of her colleague's suggestions in her assessment decisions but stated that she would like to observe these strategies in action to better understand how to implement them.

- Céline's experience with the student booklets was generally positive but she was very uneasy with the activities where students were reading each other's texts. She felt that some stronger writers were openly mocking the quality of weaker texts, but she did mention that the written feedback was always respectful of the student and even useful in some cases. She plans to use the booklets again in subsequent years and may follow through with integrating reciprocal assessment and peer feedback activities to support writing depending on the classroom culture.

## 20.5 Conclusion

One issue in the successful implementation of FA in the classroom has to do with teachers' representation of what FA involves in practice. Mireille and Céline's pathways through the three phases of the research were very different as a result of their distinct initial representations of FA.

While Mireille started out with a representation of FA based largely on teacher regulation and feedback, Céline's representation of FA was grounded in getting students to understand her expectations related to the summative evaluation. Although it is possible to observe changes in FA practice throughout the research for both teachers, the extent of these changes was not the same in each case.

To a certain extent, behavior changes before beliefs (Fullan 2001). Mireille's capacity to reconsider peer assessment as a viable form of regulation in writing after many failed attempts may have been the result of using the student booklets. Learning about how to give feedback is much the same as learning about anything else and giving students the opportunity to learn about the assessment process through co-constructing criteria and using standards is necessary before peer feedback can be put to good use. Only after having a positive experience with peer feedback using the booklets could she consider widening the scope of her representation of FA and attempt to integrate co-regulated activities in her practice.

Many changes she integrated in her practice involved replicating the activities from the student booklet, but she also adapted some activities to fit her specific needs and those of her class. For example, since she had never co-constructed criteria with students, she decided that instead of doing this at the onset of a unit, she would have students write a first draft of a text, develop exemplars from their texts, and have them differentiate among the exemplars using their reasoning to build the assessment criteria with them. She felt this was a better fit for her current level of proficiency and her students' skills. Her participation in the research challenged her initial conception of FA and helped her to consider other ways to support student learning beyond teacher regulation.

For the students in Mireille's class, participating in co-regulated activities such as the co-construction of criteria and peer feedback throughout phase 3 of the research gave them a better understanding of how the assessment process works,

which, in turn, gave them more control over the outcome. This is especially important for weaker students who tend to believe that good marks are the result of good luck rather than targeted work.

Céline's representation of FA remained relatively unchanged throughout the research. Whenever she collected information on students in writing it was to determine to what degree they met the curriculum expectations and to inform parents about how their child was doing rather than to inform her teaching practice and support learning. Although she attempted to give students the opportunity to improve their text using her feedback, she was unable to consider regulations as an integral part of teaching and learning. Her participation in the research never really challenged her conception of FA or brought about questions about the efficacy of her current teaching practice to support learning.

Educational contexts have a long history of superficial implementation of effective practice supported by research and mandated by policy. A number of studies (e.g., Weiss et al. 2003) show evidence that FA is either absent or superficial in most classrooms, which undermines its potential for promoting improved teaching and student learning. The case study of Céline is not unique. Overwhelmed by summative demands, teachers tend to comply with what is specifically prescribed by policy (e.g., number of summative assessment) while cutting back on what is not so easily measurable (e.g., providing targeted instruction to help students become learning resources for one another). There are additional reasons for superficial attempts to implement FA—beliefs about the nature of teaching and learning, lack of deep understanding of what FA involves in practice, limited opportunities for sustained professional learning—however, the overwhelming demands of summative assessment are sufficient to discourage most teachers from delving further into their practice.

For the students in Céline's class, the use of the student booklets gave them a better idea of what quality work looks like in writing but without further sustained co-regulated activities in other writing contexts; the underlying message still remained that there are few opportunities to improve learning and little hope of developing writing skills. Throughout the three phases of the research, they experienced assessment primarily through summative requirements and had few new formative experiences.

For Céline, the use of the student booklets to support peer feedback and student self-assessment may not have been an appropriate fit for her professional learning needs. It is as though FA evolves on a continuum where teachers must first consider their responsibility to support learning through teacher regulations before being able to consider other forms of regulation such as co-regulation. A unit of study where teacher regulations—proactive, interactive, and retroactive—are embedded at key points in a lesson may have better supported a change in Céline's beliefs about teaching and learning, and a deeper understanding of FA. Furthermore, she may have benefitted from observing a colleague who was effectively implementing teacher regulations in his or her classroom and being observed by that colleague to get feedback (Elmore 2004). Much in the same way that co-regulation can support

student learning and students as learners, professional development might be better conceived as co-regulation where teachers work jointly to support one another in professional learning.

Formative assessment will not meet its promise to support learning until teachers engage in different types of regulations to support learning. Although widening the scope of the types of regulation is an objective, the goal is to ensure that teachers determine the best course of action for each student's learning in a given classroom context while considering the full array of possible regulations. If interaction with peers can be an effective way to help students become self-regulated learners, it could also be an effective way to hone individual teachers' FA skills and provide them the support they need to make better assessment decisions in their classrooms. In this respect, implementing formative assessment changes the culture of the classroom where students come to understand that there are in fact opportunities for them to improve learning and develop as a learner.

## References

- Allal, L. (1988). Vers un élargissement de la pédagogie de maîtrise: processus de régulation interactive, rétroactive et proactive. In M. Huberman (Ed.), *Assurer la réussite des apprentissages scolaires. Les propositions de la pédagogie de la maîtrise* (pp. 86–126). Paris: Delachaux et Niestlé.
- Allal, L. (2010). Assessment and the regulation of learning. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (Vol. 3, pp. 348–352). Oxford: Elsevier.
- Allal, L. (2011). Pedagogy, didactics and the co-regulation of learning: A perspective from the French-language world of educational research. *Research Papers in Education*, 26(3), 329–336.
- Allal, L., & Mottier Lopez, L. (2005). Formative assessment of learning: A review of publications in French. In OCDE (Ed.), *Formative assessment: Improving learning in secondary classrooms* (p. 241–264). Paris: OCDE.
- Assessment Reform Group. (2002). *Assessment for learning: 10 principles. Research-based principles to guide classroom practice*. Cambridge: University of Cambridge School of Education. [https://assessmentreformgroup.files.wordpress.com/2012/01/10principles\\_english.pdf](https://assessmentreformgroup.files.wordpress.com/2012/01/10principles_english.pdf). Accessed July 28, 2015.
- Baumeister, R. F., Heatherton, T. F., & Tice, D. M. (1994). *Losing control: How and why people fail at self-regulation*. San Diego: Academic.
- Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and Instruction*, 7, 161–186.
- Elmore, R. (2004). *School reform from the inside out*. Cambridge: Harvard University Press.
- Fullan, M. (2001). *Leading in a culture of change*. San Francisco: Jossey-Bass.
- Hadwin, A. F., & Oshige, M. (2011). Self-regulation, co-regulation, and socially-shared regulation: Exploring perspectives of social in self-regulated learning theory. *Teachers College Record*, 113(2), 240–264.
- Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment. *Educational Measurement: Issues and Practice*, 28(3), 24–31.
- Laveault, D. (2000). *La régulation des apprentissages et la motivation scolaire*. Québec: Ministère de l'Éducation du Québec.
- Laveault, D. (2006, November 2). Putting the right accent on learning: The role of regulation in formative assessment. *The Orbit*, 36, 7–9.

- Laveault, D. (2007). De la régulation au réglage: étude des dispositifs d'évaluation favorisant l'autorégulation des apprentissages. In L. Allal & L. Mottier Lopez (Eds.), *Régulation des apprentissages en situation scolaire et en formation* (pp. 207–234). Bruxelles: De Boeck.
- Laveault, D., Bourgeois, L., & Rabideau, C. (2013). *Je rédige un résumé et je donne une rétroaction efficace!* Unpublished student booklet and teacher guide. Faculty of Education, University of Ottawa, Ontario, Canada.
- Laveault, D. (2014). The power of learning-centered task design: An exercise in the application of the variation principle. In C. Wyatt-Smith, V. Klenowski, & R. Colbert (Eds.), *Designing assessment for quality learning* (pp. 109–121). Dordrecht: Springer.
- Meirieu, P. (2006). *Faire l'école, faire la classe*. Paris: ESF Éditeur.
- Ontario Ministry of Education. (2010). *Growing success: Assessment, evaluation, and reporting in Ontario's schools, grades 1–12*. Toronto: Queen's Printer for Ontario. <http://www.edu.gov.on.ca/eng/policyfunding/growSuccess.pdf>. Accessed July 29, 2015.
- Perrenoud, P. (1998). From formative evaluation to a controlled regulation of learning processes: Towards a wider conceptual field. *Assessment in Education*, 5(1), 85–102.
- Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144.
- Scallon, G. (2000). *L'évaluation formative*. Saint-Laurent: Éditions du Renouveau Pédagogique.
- Stiggins, R. J., Arter, J. A., Chappuis, J., & Chappuis, S. (2004). *Classroom assessment for student learning*. Portland: Assessment Training Institute.
- Weiss, I. R., Pasley, J. D., Smith, P. S., Banilower, E. R., & Heck, D. J. (2003). *Looking inside the classroom: A study of K–12 mathematics and science education in the United States*. Chapel Hill: Horizon Research Inc.
- William, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37, 3–14.