Accountability and Educational Improvement

David Godfrey *Editor*

School Peer Review for Educational Improvement and Accountability

Theory, Practice and Policy Implications

Springer

Accountability and Educational Improvement

Series Editors
Melanie Ehren, UCL Institute of Education, University College London, London, UK
Katharina Maag Merki, Institut für Erziehungswissenschaft, Universität Zürich, Zürich, Switzerland

This book series intends to bring together an array of theoretical and empirical research into accountability systems, external and internal evaluation, educational improvement, and their impact on teaching, learning and achievement of the students in a multilevel context. The series will address how different types of accountability and evaluation systems (e.g. school inspections, test-based accountability, merit pay, internal evaluations, peer review) have an impact (both intended and unintended) on educational improvement, particularly of education systems, schools, and teachers. The series addresses questions on the impact of different types of evaluation and accountability systems on equal opportunities in education, school improvement and teaching and learning in the classroom, and methods to study these questions. Theoretical foundations of educational improvement, accountability and evaluation systems will specifically be addressed (e.g. principal-agent theory, rational choice theory, cybernetics, goal setting theory, institutionalisation) to enhance our understanding of the mechanisms and processes underlying improvement through different types of (both external and internal) evaluation and accountability systems, and the context in which different types of evaluation are effective. These topics will be relevant for researchers studying the effects of such systems as well as for both practitioners and policy-makers who are in charge of the design of evaluation systems.

More information about this series at http://www.springer.com/series/13537

David Godfrey
Editor

School Peer Review for Educational Improvement and Accountability

Theory, Practice and Policy Implications

Springer

Editor
David Godfrey
UCL Institute of Education
London, UK

ISSN 2509-3320 ISSN 2509-3339 (electronic)
Accountability and Educational Improvement
ISBN 978-3-030-48129-2 ISBN 978-3-030-48130-8 (eBook)
https://doi.org/10.1007/978-3-030-48130-8

© Springer Nature Switzerland AG 2020
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, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

Foreword

A report commissioned by the Department for Education (DfE) in England (Greatbatch and Tate, 2019) concluded that ‘there is a lack of comparative research to use as a basis for making systematic evidence-based judgements on which approaches to school improvement work best and in what circumstances’ (p8). The authors also note that ‘regardless of whether school improvement systems are based on school inspections or self-evaluation by schools, all the (high performing) countries considered in this review place a strong emphasis on school-to-school collaboration and peer-to-peer support, although the mechanisms through which this is organised vary’ (p8). External evaluation is increasingly seen as limited as an effective means of securing school improvement. In England, The Office for Standards in Education Children’s Services and Skills’ (Ofsted) 2019 annual Teacher Attitude survey found that the proportion of teachers who regard the inspectorate as ‘a force for improvement in education’ had declined to only 20% agreeing with over a half (56%) disagreeing. Just over one-quarter (27%) said inspections helped individual schools improve, compared with 31% in the previous year (Ofsted 2019). How schools improve and sustain any improvements remains an on-going question.

It appears as though few countries have tried peer review or supported school self-evaluation as a system-wide strategy for school improvement. In *‘School peer review for educational improvement and accountability’* David Godfrey has done an excellent job in bringing together a collection of largely research-based accounts from around the world of how schools are increasingly using systems of self-evaluation involving professional colleagues or peers in an attempt to achieve school improvement. The book’s 13 chapters consider peer review models in six jurisdictions – England (4), Wales, Australia (2), Bulgaria, the Czech Republic and Chile – along with their specific aims, their benefits (potential and real) and the challenges of building and sustaining peer-review networks. Peer review with its professional dialogue about what is considered high quality education is seen as empowering, placing the power to improve and broaden education outcomes back into the hands of the profession rather than those of inspectors, external evaluators and policy makers. Peer reviews can provide feedback and critical friendship whilst reviewing the school’s self-evaluation and, perhaps most importantly, supporting the school in its resulting improvement efforts. The methods used may be similar to inspectors but with peer reviews the stakes are much lower, and any resulting reports are for internal use only and not made public.

A key task for all leaders is the creation of conditions that foster the continuous improvement of their organisations, their systems, processes and outcomes. Systematic and supported school self-evaluation is an important component of quality assurance. It is not new however and has a long history – in England, for example, in 1979 a local authority, Oxfordshire, introduced a process whereby schools were required to give a four-yearly account of their activities involving a self-chosen peer headteacher with outcomes reported to the school's governing body and a panel of the authority's Education Committee. About the same time the Inner London Education Authority (ILEA) introduced a similar process – Keeping the School Under Review – and the School Council's 'Guidelines for Review and Internal Development' (more commonly known as GRIDs), another form of supported school self-evaluation, was piloted in a number of municipalities in England and Wales (McMahon et al. 1984).

Evaluating the performance of schools, whether by internal mechanisms or externally through an inspection process, has been the subject of research and debate for many years. For example, Peter Mortimore, a leading educational researcher, noted in the late 1990s that the school improvement debate was between two opposing groups: the doves and the hawks. For him:

> The doves argue that unless schools are able to do things for themselves then any change is likely to remain superficial. For change to be successful and improvement to be embedded, there is a clear need for 'ownership' on the part of those responsible for 'delivering the innovation' or raising standards. The hawks, on the other hand, tend to perceive self-review or evaluation as an easy option and soft-centred, and argue that without the hard edge that external probing (supposedly) brings to a school, difficult questions and judgements will invariably be shirked. (Earley 1998, p168)

Much more recently an Organisation for Economic Co-operation and Development (OECD) report (2013) cites weaknesses in relying on either approach alone, since external evaluation can lead to game playing and internal evaluation can be subject to 'self-delusion'. The reality of course is that a combination of the two is required for successful change to occur and that both internal and external school evaluation are needed for continuous improvement. Each on its own is of limited value and neither is sufficient to bring about sustained improvement in schools. Neither a top-down, nor a bottom-up approach on its own is sufficient; self-evaluation is seen as a necessary but insufficient ingredient to stimulate school improvement. Significantly, school self-evaluation and peer review can act as an important antidote to the unhealthy dominance that the inspection discourse can play in education systems like England's (MacBeath 2006).

A related on-going debate has centred on the balance between evaluation for accountability and evaluation for improvement. For some, there is a growing tendency to use school self-evaluation for accountability purposes which may conflict with its original aims of school improvement. Peer reviews can be more development oriented (formative evaluation to 'improve') or more accountability oriented (summative evaluation to ‘prove’), although their intention is usually to be more the former. External evaluation criteria may be used when undertaking peer reviews but resulting reports are not made public (and no overall ‘grade’ awarded) which allows for greater honesty about a school’s weaknesses, while ‘inviting advice on how to improve; accounting for the success of the school as well as its areas for development and opening up collaborative and supportive dialogue between members of the reviewing team’ (Chap. 9) to the benefit of both parties – the reviewed and the reviewers.

In an era of global educational reform where high stakes accountability systems are increasingly found alongside greater school autonomy, and with responsibility for school improvement shifting from local government to schools themselves (the self-improving school system), peer-review processes, the subject of this excellent edited collection, have developed as potentially an important means to improve student outcomes and raise standards. Peer reviews or supported self-evaluation are linked more closely perhaps with high performing school systems which are centred on trust-based responsibility as opposed to those jurisdictions which rely more on external high stakes inspection and test-based accountability. However, it is not uncommon in ‘high-autonomy-high accountability’ systems (Earley and Greany, 2017) for peer reviews to become little more than schools evaluating themselves to ensure that they are ‘inspection ready’ and able to meet the inspectors’ criteria for what is said to make a ‘good school’. Is it therefore the case that peer reviews are more often about ‘self-inspection’ (Ferguson et al. 2000; MacBeath 2006) than school development – to prove rather than improve? Or do peer reviews provide an excellent platform to enable schools to free themselves from the constraints of external inspection and evaluate themselves according to their own criteria, address their own areas for development and support the sharing of effective practice within and between schools? Both sets of arguments and related issues are well presented in this volume.

Peer reviews have enormous potential to generate trust, reassert professionalism and reclaim the educational agenda from external inspection which with its common framework is said to have little scope to consider the specific aims of each school. As Taylor (2019) has asked ‘Is a common inspection framework set up to inspect one school that has resilience, perseverance and discovery amongst its aims and another that has respect, imagination and excellence?’ adding that ‘these all mean different things to different people and will be interpreted by inspectors in a diverse manner’ (p1). Peer reviews, in theory if not always in practice, enable schools to reassert their priorities and focus on the perceived needs of the children and their communities. Schools are able to select their own areas of focus or they are offered a framework to derive their own specific enquiry questions for peer review.

As Greany (Chap. 4) asks ‘do they serve to reinforce the external accountability system and quality metrics, in the process making schools more homogenous’, or do peer reviews ‘offer a means for schools to take ownership of what is meant by “quality”, enabling diverse, innovative responses to contemporary challenges’? Perhaps rather depressingly, it is reported that only a minority of the interviewees in his research saw peer review as providing ‘a space “outside” the accountability system in which to be honest without fear of reprimand' (Greany and Higham 2018, p32). This also raises interesting questions about whether peer review can build genuine lateral accountability and professional collaboration when incentives and high stakes are organised around individual school performance and inspection grades. It is argued that high stakes inspection encourages a focus on individual schools instead of supporting them in developing collective capacity and benefiting from being part of a network. Thus even in those education systems where peer review is becoming commonplace 'the hierarchical and competitive environment may either affect (its) growth, or inhibit its potential for genuine school development' (Chap. 4). Collaboration and competition are uncomfortable bedfellows with the latter often preventing the development of trust.

Not enough is yet known about the essential characteristics of effective peer review and the conditions in which it has an impact. Although admitting to a paucity of research into peer review processes this edited collection helpfully brings together the existing evidence base from across the developed world and attempts to address the questions raised above and many more. The contributors to this volume vary in their interpretations of the function and merits of peer review – is it more about self-inspection/self-policing or appreciative inquiry/emancipatory evaluation?

It is admitted that peer review is still in its infancy and that its power and potential is still emergent. The Education Development Trust (Chap. 9), a major player in the field, argue that 'as schools become more autonomous and more accountable, peer review creates a climate and a culture where connected autonomy and trust-based accountability can grow' stating that in its view 'done robustly and rigorously, peer review forms the backbone of trust based and lateral accountability that can co-exist with top down and regulatory forms of accountability'. It also argues that there is 'a growing number of leaders prepared to invest in reciprocal peer review because they believe it's the right thing to do and it gets results. They want to reclaim what it means to be a great school, and to have the necessary conversations with each other about what needs to improve' (Chap. 9). These are interesting observations which hopefully the on-going evaluation of the School Partnership programme, conducted by Anders, Godfrey and colleagues, will provide some answers.

Peer review makes a reality of collective moral purpose. As the EDT states, school leaders 'model this through being willing to hold themselves and each other to account for improvement, through being ready to make their best practice available to each other across the cluster and through their willingness to tackle issues of collective importance that will ensure that the greatest number of children and young people benefit'. However, is it more the case that many school leaders have become the 'doers' of the bidding of others rather than 'playing a lead role in shaping school leadership professionalism and education more broadly for the twenty-first century' (Cranston 2013, cited in Earley 2013, p166)? Can it help to replace the constraints of accountability and replace it with 'a new liberating professionalism for school leaders framed around notions of professional responsibility' to 'position themselves as proactive reflective leadership professionals, not reactive managers' op cit, p165)? Peer review can help shape education systems rather than simply reflect policy directives. As the editor himself also notes ‘Those leaders participating in peer reviews often cite the moral dimension of improving all children’s education, including those in partner schools’ (Chap. 13).

This is a passionate plea for the merits of peer review and an antidote to the more critical perspectives which see peer review as simply another, perhaps more sophisticated, form of performativity and surveillance. Whatever ones belief about the merits or otherwise of peer review this edited collection is a timely and important contribution to that debate.

UCL Institute of Education Peter Earley
Autumn 2019

Professor Peter Earley holds the Chair of Education Leadership and Management at the London Centre for Leadership in Learning, Department of Learning and Leadership, UCL Institute of Education, University College London.His central research interests are leadership, school improvement, professional development, inspection, self-evaluation and school governance. Recent externally funded research and evaluation projects include: leadership development for head teachers (2016); new paths or routes to headship (2015); effective head teacher performance management (2014); the changing landscape of educational leadership in England (2012); and the experiences of new head teachers (2011).He has published widely and recent publications include: *Exploring the School Leadership Landscape: Changing demands, changing realities* (Bloomsbury, 2013), *Accelerated Leadership Development: fast-tracking school leaders* (IOE, 2010 with Jeff Jones) and *Helping Staff Develop in Schools* (Sage, 2010 with Sara Bubb). His most recent book is: Earley, P and Greany, T (eds) *School Leadership and Education System Reform,* (Bloomsbury, 2017), London, UK

References

Earley, P. (2013). *Exploring the school leadership landscape: Changing demands, changing realities*, London: Bloomsbury.

Earley, P. (Ed.). (1998). *School improvement after inspection? School and LEA responses*. London: Sage.

Earley, P., & Greany, T. (Eds.). (2017). *School leadership and education system reform.* London: Bloomsbury.

Ferguson, N., Earley, P., Fidler, B., & Ouston, J. (2000). *Improving schools and inspection: The self-inspecting school*. London: Sage.

Greany, T., & Higham, R. (2018) *Hierarchy, markets and networks: Analysing the ‘self-improving school-led system’ agenda in England and the implications for schools.* London: IOE Press.

Greatbatch, D., & Tate, S. (2019). *School improvement systems in high performing countries*. London: DfE.

MacBeath, J. (2006). *School inspection and self-evaluation: Working with the New Relationship.* London: Routledge.

McMahon, A., Bolam, R., Abbott, R. & Holly, P. (1984*) Guidelines for review and internal development in schools: Secondary school handbook*. Harlow: Longman.

OECD. (2013). *Synergies for better learning: An international perspective on evaluation and assessment.* OECD Publishing, Paris.

Ofsted. (2019). *Teachers' awareness and perceptions of Ofsted.* Teacher Attitude Survey 2019 report. Ofsted & YouGov.

Taylor, D. (2019). It doesn't have to be like this. *School Leadership Today*, *9.6*, p1.

Acknowledgements

David would like to thank the contributors for working so hard to provide the wealth or original case studies and analysis contained within. The authors come from a range of backgrounds, not all academics, and many have fit the writing of their chapter around extremely busy day jobs. The results of this labour are excellent; uncovering a treasure trove of experiences of peer review around the globe never seen before in one volume. Particular thanks go to Professor Peter Earley and Dr Rupert Higham at UCL Institute of Education for giving their invaluable feedback on early drafts of two chapters in this book written by the editor.

Introduction

The context for this book is of a fast-changing environment of accountability that has seen rapid changes to the ways schools and networks of school are organised. In these new networks, centralised, top-down forms of school improvement and accountability are insufficient, lacking flexibility and disempowering to the professionals working in schools. Peer-driven approaches to quality assurance and improvement are thus increasingly being turned to.

School peer reviews are a form of internal school evaluation, driven by schools rather than externally imposed on them, as with school inspections. Schools collaborate with other schools in networks, collect data through self-evaluation and in school review visits, and provide feedback, challenge and support to each other. Many school leaders are turning to peer support and challenge in a context of sometimes punitive or narrow external accountability. The book looks at the way in which this emerging form of internal evaluation interacts, interrupts or even mimics external evaluations. The impact of peer reviews is also discussed at length, and the factors that mediate these effects.

Little so far has been written about school peer review from an international perspective. While many of the largest and more established peer-review models and organisations exist in England (e.g. Challenge Partners, Education Development Trust), with many new models emerging, particularly over the last decade (e.g. National Association of Head Teachers, Research-Informed Peer Review), the book also makes visible new examples from 5 other countries: Wales, Australia, Bulgaria, The Czech Republic and Chile. This includes detailed description and case studies of 10 different models of peer review.

Underlying the rationale for school peer review are several fields of research: external and internal school evaluation, collaborative enquiry, critical friendship, joint practice development and evaluation theory. Many of the chapters provide case studies with new empirical evidence; others provide detailed explanation of the theories of action underlying the model in question and explain how peer reviews work within a wider network to improve schools. Despite the increased use of school peer review in system reform and school improvement, very little research has been conducted on this phenomenon.

The book has three main aims:

I. To provide a thorough conceptual and theoretical description and analysis of school peer review in relation to evaluation theory
II. To explain some of the system-related factors that affect the introduction, growth, effectiveness and sustainability of peer review
III. To explore peer review and its impact in a range of case studies from different countries

The book is written primarily to academics working in the field of school leadership, educational evaluation and accountability, school improvement or research into school research- and data-use. In addition, the book will appeal to those working at the level of executive leadership in school networks, NGOs and in other government organisations with a responsibility to promote school collaboration and improvement. Students studying for MAs or PhDs in the areas of school evaluation, accountability, networked learning and so on will also find this book valuable.

I hope you agree that this collection of case studies and analysis from such a wide range of distinguished authors will add significantly to the literature on this topic.

Chapter Outline

The beginnings and growth of school peer review are described in Part I. Chapter 1 provides an overview of research into external and internal school evaluation, framing the more recent growth of peer review. Peer review is defined and the existing research on the topic is analysed.

Part II focuses on how peer review has been brought in to complement school systems, in particular its accountability and school improvement strategies. Chapter 2 looks at how Queensland cycle of external reviews have introduced peer review as an additional, and complementary way of working. Chapter 3 shows how a ‘layering’ approach can be used to introduce peer reviews alongside existing accountability structures and processes in the Welsh context.

Part III addresses the widespread use of peer reviews in the context of high stakes’ accountability and there intended and unintended consequences. Both chapters focus on England. Chapter 4 looks at analyses peer reviews between schools in England through the lens of isomorphism, addressing the issue of whether schools use peer review for improvement purposes or more as a form of self-policing. Chapter 5 looks at how Ofsted inspections in England affect the conduct and impact of a peer review scheme introduced by the National Association of Head teachers.

Part IV looks at the introduction of peer reviews in school systems that have not had a prior history of such practice, or indeed an established culture of school self-evaluation.

Chapter 6 examines an ambitious project as part of EU funded research, to collaborate between researchers, a group of schools and inspectors to use a polycentric approach to evaluating a network. Chapter 7 explores the challenges of introducing widespread use of peer review in a country where self-evaluation is used inconsistently and without a supportive national infrastructure.

Part V looks at peer reviews that form part of large school improvement partnerships. Chapter 8 looks at a pioneer organisation in England, Challenge Partners, and how this has been set up to encourage system-wide improvement. Chapter 9 looks at the largest such programme in England, the Schools Partnership Programme. In both of these chapters the theory of action and evaluation principles are made visible; showing how they are designed to lead to school improvement and what Matthews and Headon (2015) describe as ‘multiple gains’.

Part VI Describes and evaluates participatory evaluation approaches to peer review, i.e. those more clearly guided by university staff and/or in the involvement of stakeholders in the construction of the evaluation approach. Chapter 10 explores case studies of principals in New South Wales in Australia who used an empowerment evaluation approach to lead school improvement with partner schools. Chapter 11 explores the approach of an innovative model of peer review, involving direct engagement with published academic research and facilitated by staff from University College London (UCL) working alongside practitioners in school clusters. Chapter 12 looks at the challenges of changing school leaders’ conversations towards professional development and learning and away from supervision and control in the context of a peer review project in Chile.

Finally, the last part attempts a much-needed synthesis of the research so far. Chapter 13 conducts a detailed comparative analysis of the above models of peer review in relation to evaluation theory before discussing a range of issues that have emerged from the research, such as the benefits of peer review as a form of leadership development and the issue of trust.

David Godfrey

Reference

Matthews, A. P., & Headon, M. (2015). *Multiple gains: An independent evaluation of Challenge Partners’ peer reviews of schools*. Institute of Education Press.

Abbreviations

Academies	State-funded schools in England which are directly funded by the Department for Education and independent of local authority control
CAs	Challenge Advisors. These come from consortia (see below) in Wales and conduct annual reviews of schools, placing them into one of four categories according to requirements for support (green, yellow, amber or red)
CPD	Continuing Professional Development. In UK, this is the term used for on-going professional development in schools.
CSC	Central South Consortium is one of the four regional consortia in Wales that is responsible for school improvement and professional learning
CSI	Czech School Inspectorate, consisting of headquarters based in Prague and 14 regional inspectorates. Inspects *Pre-school, Basic, Secondary, Tertiary Professional and Other Education*
DfE	The Department for Education is responsible for children's services and education, including early years, schools, higher and further education policy, apprenticeships and wider skills in England
EE	External Evaluation is a process of evaluation or accountability imposed by an outside body, especially those conducting on behalf of local or central government, such as inspectorates
Estyn	The national inspectorate for education and training in Wales, independent of, but funded by, the Welsh National Assembly
EVOS	Programme for Quality Evaluation in Vocational Schools in Czech Republic (1995–2000)

IE	Internal Evaluation is a process of evaluation or accountability driven by the school or its network rather than from outside
Instead	The name of the peer review programme promoted by the National Association of Head Teachers (see above). The name plays on ‘Ofsted’, i.e. providing an alternative
(Institutional) Isomorphism	The process by which organisations become similar in structure to each other
LA	Local Authority (usually referring to Local Education Authority (LEA) in the context of this book). These are the local councils in England and Wales that are responsible for education within their jurisdiction
London Challenge	The London Challenge school improvement programme ran from 2002 to 2011 and was designed to raise the performance of schools in London through a series of connected initiatives
NAHT	The National Association of Head Teachers. In England, this is a trade union that represents head teachers, deputy and assistant heads, school business leaders, special educational needs coordinators, virtual school heads and leaders of outdoor education centres
NIE	Bulgaria: National inspectorate of education. Conducts inspections on behalf of the state but retains independence from the Ministry of Education
NSIT	National School Improvement Tool. A school improvement framework used in external evaluations in Australia. Developed by the Australian Council for Educational Research
MATs	Multi-Academy Trusts. In England, these are groups of independent, publicly funded schools with a central administrative body called a trust
Mocksted	England: A practice or rehearsal for an Ofsted inspection
OECD	The Organisation for Economic Co-operation and Development
Ofsted	The Office for Standards in Education, Children’s Services and Skills is a non-ministerial department of the UK government, reporting to Parliament. Ofsted is responsible for inspecting a range of educational institutions, including state schools and some independent (i.e. privately funded) schools
Peer review	The evaluation of work by one or more people with similar competences as the producers of the work

	(peers). Synonyms include: collective review, peer evaluation, peer enquiry, collaborative peer enquiry
QAR	Quality Assurance Review. The name given by Challenge Partners to the process of peer evaluating and visiting schools (the peer-review visit)
RDE	Bulgaria: Regional Departments of Education (former Regional Inspectorates of Education), responsible for controlling and supporting schools within their regions and for organising the national exams
SIGs	School Improvement Groups, formed as part of the Central South Wales Challenge
SIP	England: School Improvement Partner. A senior professional designated by the local authority to review a school's performance and provide professional challenge and support
SIU	School Improvement Unit. Located in the Queensland Department of Education and responsible for implementing cyclical school (or system) reviews
Teaching Schools	England: These schools are designated by the Department for Education and have additional responsibilities for school to school improvement, teacher training, leadership development and promoting evidence informed practice
TSAs	England: Teaching Schools Alliances. Groups of schools with a teaching school (above) at its hub. These work together for school improvement and training purposes
VET	Vocational education and training

Contents

Part I The Emergence and Growth of School Peer Review

1 From External Evaluation, to School Self-evaluation, to Peer Review 3
David Godfrey

Part II Supplementing the Regional or National Accountability System

2 Peer Reviews as a Complement to System Reviews in Queensland 27
Chris Diamond and Anetta Kowalkiewicz

3 'Layering' Peer Enquiry as a System Change Strategy: Some Lessons from Wales 51
Mark Hadfield and Mel Ainscow

Part III Peer Reviews in the Context of High Stakes' Accountability. Intended and Unintended Consequences

4 Self-Policing or Self-Improving?: Analysing Peer Reviews Between Schools in England Through the Lens of Isomorphism 71
Toby Greany

5 Case Study of a Cluster in the National Association of Head Teachers' 'Instead' Peer Review in England 95
David Godfrey and Melanie Ehren

Part IV Peer Review in Unfamiliar National Contexts: Successes and Challenges

6 Peer Review Network of Schools – Lessons from Innovative Practice in Bulgaria 119
Rossitsa Simeonova and Yonka Parvanova

7 Peer Review in Czech Education: A Recognized but Somewhat Neglected Tool for School Development 139
Stanislav Michek, Martin Chvál, and Milan Pol

Part V Peer Review Within School Improvement Partnerships

8 The Development of a System Model of Peer Review and School Improvement: Challenge Partners 159
George Berwick

9 Education Development Trust's Schools Partnership Programme: A Collaborative School Improvement Movement 181
Andrew Ettinger, John Cronin, and Maggie Farrar

Part VI Participatory Evaluation Approaches to Peer Review

10 Empowering Principals in Peer Review: The Value of an Empowerment Evaluation Approach for Educational Improvement 201
Kerrie Ikin

11 Research-Informed Peer Review 223
David Godfrey and Karen Spence-Thomas

12 Changing School Leaders' Conversations about Teaching and Learning through a Peer Review Process Implemented in Nine Public Schools in Chile 245
Mónica Cortez, Fabián Campos, Carmen Montecinos, Jorge Rojas, Marcela Peña, Jorge Gajardo, Jorge Ulloa, and Charles Albornoz

Part VII Synthesis and Discussion

13 Evaluation Theory and Peer Review. Practice, Policy and Research Implications 269
David Godfrey

About the Editor

David Godfrey is an Associate Professor in Education, Leadership and Management at UCL Institute of Education in London and the programme leader for the MA Educational Leadership. He was co-director of the Centre for Educational Evaluation and Accountability until 2018 and was a lead inspector for the Independent Schools Inspectorate. An advocate of research-informed practice in education, his projects and publications include research-engaged schools, school peer review, inspection systems and lesson study. In July 2017, David was acknowledged in the Oxford Review of Education as one of the best new educational researchers in the UK.

Part I

The Emergence and Growth of School Peer Review

Chapter 1
From External Evaluation, to School Self-evaluation, to Peer Review

David Godfrey

Abstract Most modern systems of school education around the world now have highly developed evaluation processes. Following quickly in the wake of external evaluation policies have followed calls for schools to develop their own capacities for self-review. An OECD report (2013) describes a number of ways in which developing school evaluation capacity should be a priority for school improvement. Among the report's suggestions are promoting peer learning among schools (pp. 469–470). This chapter describes the research on external evaluation, internal evaluation and the relationship between these two. It then moves on to define and describes peer review and chart its growth. The accountability dimension of peer review is also explored. Finally the existing evidence on peer review is presented from the vocational and schools sectors to show key findings so far, and key conditions for effective peer review are described.

1.1 Introduction

Most modern systems of school education around the world now have highly developed evaluation processes. Many countries have introduced national agencies tasked with measuring indicators of educational quality, such as school inspections or test-based accountability. Following quickly in the wake of external evaluation (EE) policies have followed calls for schools to develop their own capacities for self-review.

An OECD report (2013) describes a number of ways in which developing school evaluation capacity should be a priority for school improvement. Among the report's suggestions are:

- Strengthening school principals' capacity to stimulate an effective school self-evaluation culture

---

D. Godfrey (✉)
UCL Institute of Education, London, UK
e-mail: david.godfrey@ucl.ac.uk

© Springer Nature Switzerland AG 2020

D. Godfrey (ed.), *School Peer Review for Educational Improvement and Accountability*, Accountability and Educational Improvement,
https://doi.org/10.1007/978-3-030-48130-8_1

- Promoting the engagement of all school staff and students in school self-evaluation and
- ***Promoting peer learning among schools*** [emphasis added] (OECD 2013, pp. 469–470)

On this last point, the OECD report points out that peer learning is particularly useful in systems where schools have a high degree of school autonomy, in order to prevent them from forming an introspective and defensive culture. Through partnerships, groups of schools can stimulate collegial networking, peer exchange, sharing and critiquing of practice, and fostering a sense of common direction *(*ibid*,* p. 470*)*.

Partly, this movement towards peer learning can be seen in the context of school systems that have matured from ones dominated by top-down external inspection towards ones with increased professional or lateral forms of accountability. Another aspect is the growth of networked approaches to school improvement in many countries. Peer reviews can be seen as an essential part of a network's own evaluation and improvement strategy, having the potential to drive individual school improvement and also to support network level outcomes.

In England for instance, we can see how peer review fits very well into the drive towards a so-called self-improving school system (SISS). The SISS signifies a system of sustainable improvement of schools with a strong focus on bottom-up approaches and locally embedded activities. The conditions of such a system are:

- a structure of schools working in clusters and partnerships to promote improvement
- a culture of constructing and implementing local approaches for improvement; addressing topics that are relevant for a specific locality
- highly qualified people who act as system leaders in creating new knowledge, disseminating knowledge and bringing schools together in partnership work (Hargreaves 2010).

Peer review relationships offer solutions within this framework and also balance out the perception of an overbearing top down external inspection or accountability framework. This appeal to the UK Government from the General Secretary of the National Association of Head Teachers in England makes such a point (see also Chap. 5 for more on the Instead peer review model):

> Our education system has transformed itself from the era when Ofsted was first designed, and the Committee should look at the need for Ofsted to change too. Inspections should not be adversarial; they should be a constructive dialogue between the inspectorate and professionals. 80 per cent on England's schools are rated good or outstanding, so NAHT believes the time is right to see greater peer-review within the inspection framework. Our own Instead project offers one such model that would help do that; something we hope the Committee will consider in its inquiry (Russell Hobby, General Secretary NAHT Education Select Committee, 02/03/2016).

1.2 Peer Review Outside the School Sector

Peer review has been an integral part of many professional fields outside the school sector.

A basic definition of peer review is:

> the evaluation of work by one or more people with similar competences as the producers of the work (peers). It functions as a form of self-regulation by qualified members of a profession within the relevant field (Wikipedia, accessed, 6th June, 2019).

Peer review is used in accounting, law, engineering (e.g., software peer review, technical peer review), aviation, and even forest fire management (ibid). There is a long history of peer review in Higher Education (e.g. Harman 1998) and in Further Education, peer reviews have been promoted across Europe (e.g. Gutknecht-Gmeiner 2013) and largely precede their use in the mainstream school sector.[1]

Examples of peer review activities across a range of sectors includes:

- Opening up (e.g. government) policies to scrutiny of others (other country representatives). These may include visits by teams to the host country to evaluate the success of an initiative
- Local area services (e.g. Martin and Jeffes 2011) or institutional (e.g. schools/colleges) quality assurance in which teams of colleagues from equivalent services in other areas visit to identifying weaknesses, or validate existing good practice
- Submitting scholarly work (e.g. a journal article) to the (usually double-blind) review of scholars deemed to be academically qualified to make a judgement about the quality and suitability of this work
- The evaluation of professionals' practice by other professionals in the same practice, e.g. clinical peer review of heath care professionals.

Professional bodies also recommend protocols and principles for the conduct of peer review, such as those offered by the Dental Defence Union (DDU) who encourage small groups of dental professionals to work together to improve the quality of service, by reviewing aspects of practice, sharing experiences and identifying areas for change.[2]

Looking across the range of uses in different professions, a number of reasons are given for peer review:

- To evaluate the performance of professional practice, policy or initiative by those qualified in the field to do so
- To decide if work meets the necessary standards required (e.g. for publication or funding)

---

[1] One scheme in Hampshire colleges in England has been running since 1993: http://www.eqr.org.uk

[2] https://www.theddu.com/guidance-and-advice/guides/clinical-audit-peer-review-and-cpd

- To ensure that colleagues' work is of a sufficient standard and thereby protect the status of the profession and one's own risks from sanction (Edwards and Benjamin 2009)
- To share ideas, learning and plans for improvement among groups of professionals in order to improve 'client' outcomes (e.g. patients)
- To prepare for an external evaluation, inspection or audit.

The process of peer review, particularly how it is applied to scholarly uses, has also undergone much scrutiny and criticism. For example, there are concerns that blind or anonymous review does not work because the reviewers can easily guess the identities of (particularly reputable) scholars, so reviews are not really blind. Furthermore, peer review have long been criticised for the tendency toward confirmatory bias, meaning that reviewers judge the standards of scholarly work more according to established knowledge and reject work that falls outside this (e.g. Mahoney 1977). Peer review of scholarly work can also disfavour the work of minority groups or lead to nepotism (Wenneras and Wold 2001). The judgement of peer reviewers in science has also been shown to be highly unreliable meaning that the probability of receiving research grants is due to little more than chance (Cole and Simon 1981).

Criticisms of cosy relationships, lack of rigour, subjectivity and insufficient evaluation skills of peers can also be applied to school peer review, and therefore require consideration in its implementation.

1.3 Outline of the Chapter

This chapter begins by examining peer review in relation to the general research base on internal evaluation (IE). The emergence of IE in response to and in relationship with, external school evaluation and other accountability measures is examined. Internal evaluation is defined, its rationale explained and, drawing upon relevant research the positive impact and the unintended effects of IE are shown. In addition, the conditions for effective IE are described along with what the research has shown about the interaction of IE and EE.

Following this, the chapter looks at the emergence of peer review as a form of internal evaluation. Peer review in the school sector is defined and described, delineating it from other activities. This is followed by an explanation of the different accountability relationships in peer review activities compared to external evaluations and inspection. The chapter then looks at the current research base for peer review in schools, as well as a major European initiative in the Further Education sector. The main effects and impact of peer review are presented as revealed by current research. Finally, attempts to introduce standards and accreditation for peer review in the English school system are described. I outline what the research has so far shown about the effects for peer review and what we know so far about the conditions needed for effective peer review.

1.4 From External to Internal Evaluation

Many countries have developed strong external evaluation regimes to hold their schools to account and to promote improvement. This is despite a dearth of evidence (especially outside of the UK) that inspection has an overall beneficial effect on the school system with some reported benefits balanced by unintended negative consequences, such as gaming (e.g. Nelson and Ehren 2014). The perceived imposition of external standards on the teaching profession and on school leaders, alongside the need to be prepared for external evaluations, often with high stakes, has led to a strong drive to promote school self-evaluation (SSE). Indeed, as will be discussed below, many proponents of external evaluation have suggested that the introduction of internal evaluation would follow naturally from the introduction of inspections and that these could be complementary and/or an expression of the maturity of the system (e.g. Hargreaves 2012; Barber 2004). Thus, in the late 1990s many international school systems began to focus their attention on SSE, including most of the more developed European countries, Canada, New Zealand and Australia (Hofman et al. 2009).

In the case of England, the creation of the Office for Standards in Education (Ofsted) in 1992 formed the vision of the then Conservative Government for a strong central external system to hold schools to account. This was coupled with a range of other measures that allowed parents to see public data about all schools, allowing the media to convert these into performance tables of examination results. In response to this powerful set of external accountability measures a study was commissioned by the National Union of Teachers (NUT) into self-evaluation and later published in a book called *Schools Must Speak for Themselves* (MacBeath et al. 1995). The incoming Labour Government viewed SSE very positively, and Ofsted endorsed a framework for self-evaluation inspired by the lead taken in Scotland. This was followed by the publication of a range of quality indicators and sources of data that allowed schools to compare themselves to national benchmarks, at school and individual pupil level, including value-added data (MacBeath 2005). SSE was specifically promoted at system level and with the support of the external inspectorate, with the aim of allowing schools to make targeted school improvement efforts based on comprehensive outcome data. The SSE was designed to mirror that used by the inspectors and thus form a sequential process of school improvement, initiated by the school itself.

Most OECD countries now have varying degrees of legal requirements in place for schools to conduct self-evaluation (OECD 2013). Some authors, such as MacBeath have long advocated strongly for SSE in all school systems in order to unleash systemic improvement (e.g. MacBeath 2005). Although most systems that employ self-evaluation routinely have some kind of SSE followed by school inspection, other countries such as the USA with data/test-based accountability models have shown interest in SSE approaches. This is because they focus more on the processes that drive school improvement and a wider range of educational indicators than test-based accountability alone (Ryan et al. 2013).

The notion of self-evaluation relates to a number of related concepts found not only in education but in other sectors. MacBeath (2006 p. 4) outlines these, noting that terms convey particular stances of national systems or summative/formative intentions, such as:

Audit: these suggest taking stock of resources and are summative.

Quality assurance: these are a systematic, and usually external-conducted type of audit, carried out for accountability purposes.

Self-review: these can be synonymous with self-evaluation and often used to indicate a summative over-view e.g. of a whole school rather than targeted areas.

Self-assessment: this may involve an examination of the knowledge, skills and attitudes gained by pupils and can be both summative and formative.

Inquiry or Appreciative Inquiry: used more in North America, Inquiry or AI focus on how an organisation can evaluate its strengths within its own frame of reference. This is essentially formative in nature.

Research: this is sometimes used as a synonym for inquiry (or enquiry in the UK context). This can be formative or summative, often involving a range of stakeholders including student and teacher researchers. (See ‘research-engaged schools’, for example Godfrey and Brown 2019).

Self-evaluation: this is meant to be a formative process, embedded into cyclical school practices and linked to pupil learning and achievement.

Another way of characterizing the distinction between the summative or formative purposes of these types of evaluations is made in MacBeath’s (2005) comparison of self-inspection (summative) with self-evaluation (formative). Self-inspection (a term first found in Ferguson et al. (2000)) is top-down, a one-off event providing a snapshot, is accountability focused and based on a rigid framework and pre-determined criteria. Such a process tends to be risk averse and intent on showing how the school meets its standards rather than how it could exceed them or work on its weaknesses. Self-evaluation is, by contrast, bottom-up, continuous, provides an evolving picture, is flexible, creates relevant criteria and focuses on improving teaching and learning (MacBeath 2005, p. 45). This distinction is also helpful and relevant in the consideration of peer reviews, since there is a risk for visiting reviewers of adopting an inspectorial approach with the consequence of limiting professional dialogue and opportunities for school development (see Chap. 5).

1.5 Internal Evaluation in Schools

1.5.1 Defining Internal Evaluation

The broader term internal evaluation (IE), while often used synonymously with SSE, can be used to make a clear distinction between an accountability process driven from the outside (often a government agency/inspectorate) and one driven from the inside (such as a school or school network). While staff external to the school may assist internal evaluations, these are at the behest of the school leadership and focus on areas that are pre-determined by them. Therefore, school peer reviews are best described as a form of internal evaluation, albeit some authors described them as external evaluations due to the visit of staff outside the institution (e.g. Stinton 2007; Gutknecht-Gmeiner 2013). The section below further clarifies definitions of school peer review as described in the chapters of this book, including this slightly grey area. However, to begin with, I use internal evaluation as the umbrella term for both single school self-evaluations and peer reviews conducted over multiple institutions or sites. Much of the research focuses specifically on SSE, however, given the overlap in these processes and sometimes synonymous use, lessons can be drawn from existing research in this area that apply as well to peer review. Later I focus more on the incipient research on peer review specifically, as indeed do the succeeding chapters in this book.

1.5.2 The Effects of Internal Evaluation

A systematic review of the literature on internal evaluation by Nelson et al. (2015) looked at the effects of internal evaluation in schools. The report, synthesising research from 1998 to 2015, considered empirical evidence from 20 countries. Below I summarise the findings from the synthesis in terms of: the beneficial impacts of IE and how to ensure these are maximized; the unintended effects of IE and the conditions for effective IE.

1.5.3 The Positive Impact of Internal Evaluation

The international evidence shows that school internal evaluations have led to a number of positive changes such as:

- increased reflection on school quality and intentions to improve
- greater sensitivity to areas in need of improvement, informing goals and actions for improvement planning
- the identification of professional development needs

- greater ownership of change
- increased professional learning
- revised content or organisation of the curriculum
- the provision of targeted support for groups of pupils (Nelson et al. 2015)

Student achievement is maximized when the internal evaluation is more accurate and school improvement priorities are very specific; when student under-achievement is identified and targeted for improvement and where it is used to drive professional development. Changes to teachers' practices are more likely when schools use an enquiry-based protocol to examine student data and use this to foster the acquisition of the teaching skills and knowledge needed to raise student attainment. Finally, positive outcomes from IE occur when teachers attribute student achievement to their own teaching rather than to external causes, increasing collective teacher efficacy (see Chap. 12). Internal evaluations are less successful when teachers and leaders are not supported to use or implement IE and use informal methods, when school leaders are unable to interpret data accurately, and when no time is set aside to interpret or act upon the data collected.

1.5.4 Unintended Effects of Internal Evaluation

As well as the many desirable outcomes from IE, there are also some negative consequences that occur when it is not implemented under the best conditions. The 2015 review points to examples when IE can lead to a fixation on measurement and performativity and compliance rather than improvement. An over-reliance on test-based accountability in IE can also lead to the neglect of other achievements and priorities for evaluation. Staff can also suffer from stress, anxiety, and an increased workload.

1.5.5 Conditions for Effective Internal Evaluation

Given the above potential positive effects of IE, and what is known about potential negative outcomes, the following conditions for effective IE emerge from the literature:

- School staff need to develop evaluation literacy, i.e. how to use research-related, enquiry skills
- IE needs to be properly resourced – especially using effective, validated tools for data collection and adequate time for analysis
- Leadership should focus on the development of an enquiry-oriented culture and endorse the importance of IE.
- District level (or middle-tier/network) support and guidance should be given along with expectations for IE and enable sufficient time to conduct it

- External partners can offer critical friendship and support (e.g. universities/school improvement specialist partners)
- A climate of trust needs to exist and/or be developed in the school and between partners.

The last two points taken from the Nelson et al. review are particularly relevant to peer reviews, where trust is needed to share data and work collaboratively across multiple schools and where external support is required to guide the process of reviews.

1.6 The Interaction of Internal and External Evaluation

The wider context of the accountability system also needs to be taken into account when considering internal evaluation. For instance, some systems (such as England) have very high stakes external evaluations and this may lead to the 'rehearsal' model of IE as schools feel the need to be in a state of readiness for external inspections.

Janssens and Van Amelsvoort (2008) conducted an exploratory study into the effects of School Self-Evaluation (SSE) used by eight Education Inspectorates in seven European countries: England, Scotland, the Netherlands, Northern Ireland, Denmark, Belgium and Germany (Hesse and Lower Saxony). A research team from the Netherlands inspectorate analysed documents provided to them by each inspectorate. The degree to which schools were given guidance on how to complete the SSE was looked at, as well as the position of the SSE in school inspections. The study explores the extent to which each system orients towards an accountability orientation (AO) or the improvement-orientation (IO) in their SSE.

In countries with an improvement orientation (supportive), the SSE occupies a weak to moderate position in the inspection process. Where the national system of inspection has an accountability orientation (AO), the SSE occupies a stronger position. A key aspect is the amount and type of steering given to schools about the SSE process and framework. The authors say that there is a growing tendency in Europe to use SSE for accountability purposes and that this may conflict with the SI aims.

Some evidence shows how external school evaluations can also strengthen internal evaluations and therefore increase their potential for school improvement. Survey research in six European countries looked at direct and indirect links between the external evaluation, SSE and school improvement. Ehren and colleagues' research showed that the role of external evaluations was to set standards for what constituted a 'good school' and to sensitize stakeholders (parents, principals, students) to external evaluation reports. Both of these in turn, led to improvements in the school's self-evaluation. The external school evaluations can stimulate SSE and this can be related to specific school improvement actions (Ehren et al. 2013). These actions concern building capacity by improving teacher participation in decision-making, improving teacher co-operation and also by improving transformational leadership. These in turn relate to improvements in school effectiveness as measured by improved opportunities to learn and assessment of students and the school. These improvements in SSE, more than the acceptance of the principle of the external feedback, were found to be the key to driving improvements.

There are three overall ways to connect internal evaluation with external evaluation (Kyriakides and Campbell 2004). These are:

1. Parallel existence: The external evaluation is more concerned with accountability and internal evaluation with school improvement
2. Sequential: The school conducts its own evaluation and then an external body uses it as a basis to conduct its own or vice-versa
3. Cooperative models: Internal and external evaluators discuss and negotiate criteria. Measurement criteria are combined in a holistic evaluation, taking into account the interests of all parties.

For the first two models – parallel and sequential -, the external evaluator should be responsible for the accountability agenda. These would tend to centralise the process, towards agencies acting for the government. In the parallel model (such as in the USA), schools would tend to conduct more improvement oriented SE, while the external evaluation is more test-based and accountability driven. The two are quite distinct and can be compared. In the sequential model (such as in Hong Kong), the external evaluators can validate or challenge a school's IE. However, the process can also work the other way, with recommendations by external evaluators being fed forward into the school's IE.

Many school systems may encourage or mandate the use the external evaluation criteria in their own IE. While this may align the two concerns somewhat, the external criteria may lack the context specificity required to drive genuine school improvement. Few systems have the cooperative model which, while possibly an ideal one, may require a degree of resourcing on the part of external evaluators that is hard to achieve in practice. The area-based inspections in Northern Ireland could be considered one such example.[3] Kyriakides and Campbell (2004) argue for a maturity model, where the contribution of the school to the process may depend on how well they have performed, moving from the first two models – parallel to sequential – to the third model of cooperation.

An OECD report (2013) suggests that school systems should set up external evaluation in ways in which internal evaluation is optimised and vice-versa. The report cites weaknesses in relying on either alone, since external evaluation can lead to game playing and internal evaluation can be subject to 'self-delusion'. They suggest a number of potential ways in which external and internal evaluations can be mutually reinforcing. While schools best know their own contexts, external evaluators can provide rigour and expertise in interpretation and validation of the school's judgements. The OECD report describes a number of ways to increase this coherence at system level, including: developing agreed national criteria on school quality; developing appropriate resources for schools to use in their self-evaluations; ensuring a strong evidence base for external school evaluation and appropriate analysis tools; and ensuring transparency in external school evaluation procedures.

---

[3] http://www.schoolinspections.eu/update-after-one-year-studying-school-inspections-in-northern-ireland/

1.7 From Within School Self-evaluation to Peer Review

So far this chapter has looked at internal evaluation more broadly and its relationship with external evaluation in school improvement. Much of the research looks at within school self-evaluation, although SSE is also an integral part of a sequence of IE that can lead to peer review.

In the schools sector, the overlap with peer observations, networked learning communities, learning rounds and collaborative enquiry means that more attention is needed to adequately describe and delineate peer review activity. Thus, I outline such a definition, below:

> School peer reviews are evaluations carried out by peers of schools or parts of schools (such as departments, subject areas or year groups). Schools nominate staff to collaborate with other schools in networks, partnerships and clusters to collect and analyse data in school review visits. These visits usually build on the school's own self-evaluation in the area of focus for the review, offering validation or challenge to the school's own findings. Visiting review teams provide feedback to the school (initially verbally) and often a report is produced for the school's internal use to summarise the findings and to give recommendations. In some reviews, participatory evaluation approaches are used (Cousins and Earl 1992) in which collaboration occurs between evaluators (or researchers) and practitioners and a wide range of stakeholders are included in the evaluation process.

In most peer review programmes, there is usually some form of mutuality or reciprocity involved, in that schools may sign up to be visited but also reviewers from this school may visit another one. In some larger organisations that conduct peer review (e.g. Challenge Partners) this mutuality may be spread out among members of a larger network, so that schools can call upon peers to conduct reviews of specific areas or their own staff may be called upon to visit another on request. The mutual learning also comes from school visits, where reviewers are learning new practices and policies while they are evaluating the host schools.

School peer reviews provide feedback, critical friendship and validation (or not) of the school's self-evaluation. Peers can also support fellow schools' improvement efforts. While peer review programmes are often structured, facilitated and accredited by external agencies (e.g. EDT, Challenge Partners, UCL Institute of Education), many schools, local authorities and networks set up their own programmes and devise their own cycles and schemes for peer review. The peer review activity is likely to be supported and structured within a network (e.g. Challenge Partners, see Chap. 8), a school improvement programme (e.g. the Schools Partnership

Programme, see Chap. 9) or it can sit within an evaluation cycle of a broader and permanent alliance of schools.[4]

As with other forms of internal evaluation, peer reviews can be more improvement oriented (formative evaluation to ‘improve’) or more accountability oriented (summative evaluation to ‘prove’), although the (at least stated) intention of peer reviews is unarguably more the former. For instance, while external evaluation criteria may be used in the conduct of school peer reviews, there is no obligation to publish reports from peer reviews as their purpose is to help the school in its improvement efforts.

Peer Reviews Are Not the Same As:

- Internal school peer observations or other within-school peer coaching or learning
- Networked learning communities or research learning communities (Brown 2017), which may involve data analysis and collaborative school enquiry but do not involve evaluative school site visits
- Inspections – even in systems where these are entirely or mainly carried out by those currently serving as school leaders. This is because the accountability dimension is different, e.g. holding the school to account on behalf of the Ministry or central government.

Although usually voluntary (at least for school leaders), some systems, such as England, are beginning to embed peer review to the extent that they have become an expected part of the internal evaluation process for many schools. Some tighter networks such as Multi-Academy Trusts (MATs) in England (groups of independent publicly funded schools with a central administrative body called a trust) may mandate peer reviews for all the schools in their trust. In these cases, the peer reviews are still ‘internal evaluations’ as they are internal to the trust, albeit may be seen as external to individual schools.

1.8 The Emergence and Growth of Peer Review in School Systems

Recent incarnations of peer review programmes have emerged in response to perceptions among school leaders and others, about the dominance of external accountability. The notion of peer review suggests a relationship in which the knowledge and perspective of each party is equally recognized. Davis and White (2001) suggest that peer review in a sense was more prevalent up to the end of the 1970s, when a professional model prevailed that:

---

[4] Listen to this example of a Teaching School Alliance in England on how they incorporate peer review: https://audioboom.com/posts/6854494-matt-davis-and-marie-claire-bretherton-talk-school-to-school-collaboration

> … assumed that educational quality was best ensured by trusting teachers, advisers and others with relevant training to make decisions in the interest of pupils. In this model professionals decided on ‘good practice’ (Davis and White 2001, p. 675).

The authors give examples of how local education authority advisors and HMIs assisted in school evaluations, going on to say that the school system (in England and Wales) has subsequently become much more centralized and standardized, reducing the power of teachers and school leaders to make their own decisions. For Barber (2004), there was a shift from informed prescription in the 1990s to informed professional judgement in the 2000s. He suggested that the locus of responsibility would shift more onto teachers and school leaders in a more mature system. While not specifically promoting peer review, he said that this period would need to include sharper, more intelligent forms of accountability.

An OECD review in 2013 reported that peer review practices were emerging across several countries, including the Czech Republic, Finland, England, Sweden and provides a case study of multiple peer review practices in Belgium. The latter case, published in an earlier OECD review (Shewbridge et al. 2011), gives examples of how schools have reported increased skills in critical friendship and in self-evaluation capacity through their involvement in collaborative peer review networks.

Although reliable data elsewhere is not available, peer review in the schools sector is probably more established in England than elsewhere in the world. In a 2019 think piece, Gilbert notes that peer review is increasingly part of local area partnerships’ change strategies and school improvement work in England (Gilbert 2017). As pointed out by Greany in Chap. 4, a survey conducted in 2018 showed that nearly half of all schools had engaged in peer review in the previous year (Greany and Higham 2018). Challenge Partners was a pioneer organisation, emerging from the London Challenge programme. Running from 2002 to 2011, this programme emphasised strong system leadership and partnership working with the creation of new roles and of designated Teaching Schools (see Chap. 10).

Other big networks have subsequently formed in England, notably the Schools Partnership Programme (SPP, see Chap. 9). Recent survey data, obtained from an ongoing evaluation by the Education Endowment Foundation (EEF) of the SPP shows this trend towards growth in uptake of peer review in schools in England. Of the 339 primary schools surveyed in June 2018, a third of the sample (111 schools or 33%) said that they had been involved in a peer review programme other than SPP over the 2 years prior to the survey. Of this third, 59% said that this was a model developed by themselves in partnership with other schools. There were also a variety of other sources for peer review, including 14% were using a local authority model or one used in part of their formal school (Multi-Academy) Trust. Some also mentioned specific providers who specialised in school improvement networks revolving around peer review, for instance Challenge Partners (10%). We should bear in mind all of these schools are taking part in the School Partnership Programme (SPP) and had not done so previously as a condition of the trial (Anders et al. forthcoming). The figures from this sample of English schools show a sizeable minority of schools used to using peer review and this figure seems likely to continue to grow.

1.9 Peer Review and the Accountability Dimension

As a form of IE, peer review fits within a wider school accountability framework and may be adding additional layers of accountability to the system above and beyond what other forms of evaluation achieve. Earley and Weindling (2004) outline four accountability relationships in the school system:

- Moral accountability (to students, parents, the community)
- Professional accountability (to colleagues and others within the same profession)
- Contractual accountability (to employers or the government)
- Market accountability (to clients, to enable them to exercise choice)

In many countries, the last two forms of accountability are heavily emphasised. These forms of holding schools to account tend to be taken up by central, external evaluation agencies (such as Ofsted in England) and have a largely summative function. The first two, argues Gilbert (2012) are less emphasised and should have more attention placed on them. These are the forms of accountability that underpin most peer review activity and can be seen for example, in the pledge that schools take when working together in the Challenge Partners' programme (Matthews and Headon 2015). While the moral dimension provides the overall rationale for working together, this is achieved through the challenge and support provided by peers. This form of professional accountability can be seen as an extension of the kinds of professional responsibility seen as inherent in many professions, notable in medicine, with the so-called hippocractic oath.[5] Gilbert (2012) cites a suggestion by a college principal for teachers' standards, adopted from the medical model which would include: "protect and promote the education of students both within your school and across the schools system", and "working with other colleagues and schools/colleges in ways that best serve the interests of all students" (Gilbert 2012, p. 11). The elements of both professional and moral accountability are clearly exemplified here and peer reviews focus around both aspects when they involve teachers (and not just school leaders) in the process.

Andreas Schleicher (2018) emphasises the importance of this form of professional accountability, in which "teachers are accountable not so much to administrative authorities but primarily to their fellow teachers and school principals" (p. 116). This has also been characterised as a shift towards greater horizontal accountability, where school-to-school networks complement the vertical accountability system of assessment and Ofsted inspection (Grayson 2019, p. 25). In many countries where school networks have been emphasised, peer reviews offer a structure for increasing levels of this kind of accountability and complementing the existing vertical mechanisms.

---

[5] now updated by the World Health Association and called a pledge, see: https://www.bioedge.org/bioethics/new-hippocratic-oath-for-doctors-approved/12496

1.10 The Research Base on School Peer Review

So far there has been little published research on peer review in schools, so questions of effectiveness, particularly on mechanisms by which they might raise student outcomes, are still unanswered. The rest of this book brings together a collection of work that will add to this knowledge base, including several empirical studies. Below I outline some other examples.

One early, published research in the Netherlands studied 27 primary schools between August 2004 and June 2006. This used a research and development approach to the implementation of a peer review model that combined self-evaluation, visitation by peers and inspection (Blok et al. 2008). This model was specifically designed to balance the demands of external and internal evaluation and avoid the loss of professional voice found in a hierarchical system of evaluation with external evaluators on top and practitioners below. In this model, the schools, divided up into three regional clusters, carried out their own self-evaluation within an agreed framework. This was followed by a one-day visitation from a team of teachers and school leaders from other schools. Schools were encouraged to take part in review visits as well as to host them, in order to create a sense of mutual dialogue and learning. The final phase involved inspectors from the Dutch national inspectorate making a preparatory visit followed by a regular inspection. The former visit was designed to clarify the school's self-evaluation and the visitation report and to set the scope of the formal inspection visit.

School principals taking part in this Dutch research reported significant benefits from the visitation of critical friends, albeit finding it time consuming. They found that the process helped them to build the school's capacity to improve (Blok et al. 2008, p. 391). However, Blok and colleagues examined the quality of reports from the SSE, the visitation and the external evaluation. They concluded that the visitation reports were often low in quality or failed to provide empirically based conclusions about the school's own self-evaluation. Interestingly though, in the 24 visitation reports they analysed, eight reports considered the school SE alone, 12 considered the SE plus other questions and four considered only 'other questions'. This was despite guidance given to all schools that ONLY the school's SE should be considered for the visits. The authors concluded nevertheless, that the focus of the visits was almost certainly agreed by the visited school and the visitation team. Thus, it may be that the benefits of the visits were not explicit in the reports themselves and these may reflect an unimportant part of the process to the schools involved. Nevertheless, this research does lend support to the idea that peer evaluations of this nature are not straightforward to either conduct or report. The positive evaluations of the clusters by participating staff were also highly likely to be partly as a result of very committed schools who had self-selected their involvement in the project.

Earlier I mentioned an ongoing trial by the EEF of the Schools Partnership Programme (SPP),[6] this is the largest such trial run so far, involving well over 300 schools in a matched propensity design, comparing the improvement trajectories of similar schools over 2 years. The study will look at improvements in numeracy and literacy scores compared to the matched schools. It must be emphasised that the SPP is not only a peer review programme but an overall 'package' of school improvement driven by partnerships of schools who self-evaluate, peer review and provide ongoing support and challenge. Nevertheless, the peer reviews form a central plank of the SPP and can be seen to be the heart of what SPP aims to do in terms of increasing professional accountability and providing agency for change to the school leaders and teachers within the partnerships. A large part of this study also looks at the implementation process, in order to tease out the intermediate variables and conditions underlying this school improvement process, following SPP's own theory of action (see also Chap. 9).

In an independent evaluation of the Challenge Partners (CP) programme (Matthews and Headon 2015, see also Chap. 8) found a number of positive impacts from peer review (quality assurance (QA) visits). The authors looked at the review system, protocols, handbooks and guidance; they conducted documentary analysis of 25 QA reports and surveyed and interviewed numerous stakeholders. The surveys included 70 headteachers, 71 lead reviewers, 200 reviewers and 43 trainee reviewers. The report concluded that there were 'multiple gains' from the CP programme, not least of these was the high quality professional/leadership development gained by participants. This came about through leading and participating in reviews and also through the training and interchange of professional dialogue. Eighty-four percent of headteachers of reviewed schools felt that reviews had been very useful to the professional development of their senior leaders and over 90% of headteachers felt that they had helped in planning school improvement. The report set out a number of conditions for successful peer reviews (see below) and also recommendations for how such a programme could contribute at system level.

1.11 A European Peer Review Network

While not in the (primary/secondary) school sector, extensive work and research has been carried out in Further Education (FE), a sector similar enough form which to draw some useful lessons.

In Europe three phases work on peer review have been implemented and studied in FE and other sectors between 2004 and 2009. So far, in the three projects (the last one finishing in 2018), 38 project partners from 15 European countries have taken part (Austria, the Czech Republic, Denmark, Finland, Germany, Hungary, Italy, the

---

[6] see: https://educationendowmentfoundation.org.uk/projects-andevaluation/projects/schools-partnership-programme-spp/

Netherlands, Portugal, Romania, Switzerland, Slovenia, Spain, Turkey, and the United Kingdom). In these projects, transnational peer reviews involved assembling a team of peer reviewers from at least three different countries to come and visit and conduct the peer review at a host institution sighed up to the project (e.g. Gutknecht-Gmeiner 2013). These projects focused first on initial vocational education and training (VET) and then adult education and non-formal and informal learning. There have also been a number of innovation transfer projects, including in Lithuania, Finland, Slovenia and France.

The initial projects were inspired by higher education's use of peer reviews to evaluate institutions or departments and adapted a model to use in initial VET. The European peer review network was set up to build on these projects and to create an ongoing Europe-wide network of events and partner institutions by the Austrian government.[7] One of the outcomes was to produce tools to measure the impact of peer review in VET and also to devise a peer review assessment tool.

Looking at the impact of the 2004–2009 phases, the research found self-reported improvements in the quality areas implemented in 13 of the 14 case studies, since the peer review took place (Gutknecht-Gmeiner 2010). The report states that:

> Impacts of these improvements as measured for instance by satisfaction rates of different stakeholders (students, staff, cooperation partners like enterprises), better achievement rates of students etc., can be detected in half the cases based on quantitative survey results and indicators or similar evidence (Gutknecht-Gmeiner 2010, p. 33).

In several of the case studies, improvements were also reported in other quality areas, suggesting the knock-on effect of peer reviews beyond their initial focus. Other changes included changes to quality assurance procedures and evaluation tools as a result of their involvement (ibid, 33–34). Finally, the report notes that, "in several cases, the pilot Peer Reviews have also influenced quality management systems on the national/regional level", including in Scotland, Romania, Austria, Denmark and Finland (ibid, p. 35). The last result may reflect some of the natural growth trajectory as mentioned above, mirroring the school sector but also most likely the positive experiences of those that took part in the scheme.

The international cooperation in this project provides a novel dimension that could be usefully applied to the school sectors as well as the work done to decide on shared tools and frameworks for peer review.

1.12 Standards for Peer Review

In the above European projects, much emphasis was given to the drawing up of processes, handbooks, guidelines, tools and standards for peer review. Indeed, these have led to European wide standards and have been adopted across many countries, as mentioned above. While the adoption of peer review has lagged behind in the school sector, various organisations have sought to do the same in the English school system. So far, each organisation involved in devising and supporting peer review schemes has drawn up its own standards, however in recent years the NAHT has set up its accountability commission which includes peer review in an overall assessment of the current system of accountability in England.

---

[7] http://www.peer-review-network.eu/pages/about-us.php

The commission's concern is that the plethora of models for peer review may hide many examples that lack the sufficient rigour to be effective. Recommendation 7 of the commission suggests the need to evaluate existing peer review programmes to identify characteristics of effective practice in order to develop national accreditation arrangements (NAHT 2018, p. 6). The commission observes:

> The English education system is on a journey; too few schools currently engage in peer review, and not enough is yet known about the essential characteristics of effective review and the conditions in which it has an impact (ibid, p. 19).

The aim of the commission is to further understand the features of effective peer review and then to establish, "accreditation arrangements to oversee the expansion of suitable models that are proven to deliver tangible and sustainable benefits" (ibid, p. 19).

1.13 Conditions for Effective Peer Review

The current evidence based already suggests several conditions under which peer review may be more or less effective, some of these will be explored further in Chap. 13, which looks at how peer reviews could be scaled up into the school national evaluation system. Here I summarise some of the key learning points.

Matthews and Headon's 2015 report highlighted several conditions for effective peer reviews:

- Having well trained, skillful lead reviewers, experienced in inspection methodology[8] and a high-quality reviewing team
- Reviews that adapt to schools' needs and desired areas of focus
- A school leadership and staff predisposed to receiving feedback positively
- A willingness to meet the cost and allocate time to reviewers (although the cost compared to other training programmes is likely to be minimal)
- Ensuring that peer reviews are conducted in a developmental rather than inspectorial approach, being open and honest about weaknesses as well as strengths
- A high level of trust across partner schools and willingness to maintain and build on this partnership work
- Sustainability and moving onto multiple cycles of review may be helped by adapting approaches, including focusing on specific areas of improvement within

---

[8] this may be balanced by the danger of approaching reviews too much as if they are real inspections (see Chaps. 4 and 5). However Matthews and Headon see as advantageous that reviewers can give an assessment of where the school would stand if inspected (ibid, p. 45).

a school or by conducting whole school reviews when schools are ‘in-between’ inspections, to act as a ‘temperature gauge’ of current performance.

One of the contributors to the NAHT commission, Kate Chhatwal of Challenge Partners, comments that there are four essential ingredients of peer review programmes[9]:

Independence: the reviews should be led by people who have sufficient distance to give an honest and impartial evaluation of the school or the network. These reviewers should evaluate not according to comparisons with own school but by asking what they are doing and what impact it is having. Reviewers can also gain from the experience and take back ideas to their own setting.

Reviewing ‘with’ the school: the review is conducted alongside the school’s leaders and generates professional dialogue rather than giving a sense that they are being ‘done to’.

Revealing the skeletons: schools need to show rather than hide any difficulties they are having from the review team.

Commitment to better outcomes: these should be shared outcomes, for all schools being reviewed not just to their own setting. Schools need to commit to working before and after their annual reviews.

Although applied to the CP programme, the above comments can be applied to all school peer review programmes. To these can be added, from the experience in the European network and in the Dutch research by Blok et al. (2008):

- the requirement for ongoing training of review teams to ensure high quality reviews
- and the use of a well thought out and shared framework for evaluation
- the use of high quality (trialed) tools for evaluation

Further research on conditions for effective peer review, as well as those that compared the efficacy of various models in use worldwide, to which this volume contributes, should add to the much-needed knowledge base in this growing movement.

Acknowledgements Many thanks to Professor Peter Earley who looked at an early draft of this chapter and provided some invaluable edits and feedback.

References

Anders, J., Godfrey, D., Stoll, L., Greany, T., McGinity, R., & Munoz-Chereau, B. (forthcoming). Evaluation of the schools partnership programme. *An evaluation for the Education Endowment Foundation.*

---

[9] https://www.challengepartners.org/news/comment-power-peer-review

Barber, M. (2004). The virtue of accountability: System redesign, inspection and incentives in the era of informed professionalism. *Journal of Education, 185*, 1–38.

Blok, H., Sleegers, P., & Karsten, S. (2008). Looking for a balance between internal and external evaluation of school quality: Evaluation of the SVI model. *Journal of Education Policy, 23*(4), 379–395.

Brown, C. (2017). Research learning communities: How the RLC approach enables teachers to use research to improve their practice and the benefits for students that occur as a result. *Research for All, 1*(2), 387–405.

Cole, S., & Simon, G. A. (1981). Chance and consensus in peer review. *Science, 214*(4523), 881–886.

Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis, 14*(4), 397–418.

Davis, A., & White, J. (2001). Accountability and school inspection: In defence of audited self-review. *Journal of Philosophy of education, 35*(4), 667–681.

Earley, P., & Weindling, D. (2004). *Understanding school leadership*. London: Sage.

Edwards, M. T., & Benjamin, E. M. (2009). The process of peer review in US hospitals. *Journal of Clinical Outcomes Management, 16*(10), 461–467.

Ehren, M. C., Altrichter, H., McNamara, G., & O'Hara, J. (2013). Impact of school inspections on improvement of schools—describing assumptions on causal mechanisms in six European countries. *Educational Assessment, Evaluation and Accountability, 25*(1), 3–43.

Ferguson, N., Earley, P., Fidler, B., & Ouston, J. (2000). *Improving schools and inspection: The self-inspecting school*. Sage.

Gilbert, C. (2012). *Towards a self-improving system: The role of school accountability*. National College for School Leadership.

Gilbert, C. (2017). *Optimism of the will: The development of local area-based education partnerships*. A think-piece. LCLL publications. Accessed 11 June 2019 at: http://www.lcll.org.uk/publications.html

Godfrey, D., & Brown, C. (Eds.). (2019). *An ecosystem for research-engaged schools: Reforming education through research*. Routledge.

Grayson, H. (2019). *School accountability in England: A critique* (Headteacher Update, Vol. 1, pp. 24–25). Published Online: 10 January 2019. https://doi.org/10.12968/htup.2019.1.24.

Greany, T., & Higham, R. (2018). *Hierarchy, markets and networks: Analysing the 'self-improving school-led system' agenda in England and the implications for schools*. London: IOE Press.

Gutknecht-Gmeiner, M. (2010). *Peer review impact analysis report. Analysis of 14 transnational European peer reviews carried out in eight European countries 2006–2009*. Accessed online 12 June 2019 at: https://www.oph.fi/download/130480_Peer_Review_Impact_Analysis_Report_10_11_29_final.pdf

Gutknecht-Gmeiner, M. (2013). Peer review as external evaluation in vocational training and adult education: Definition, experiences and recommendations for use. *Edukacja ustawiczna dorosłych, 1*(80), 84–91.

Hargreaves, D. H. (2010). *Creating a self-improving school system*. Accessed online 13 June 2019 at: https://dera.ioe.ac.uk/2093/1/download%3Fid%3D133672%26filename%3Dcreating-a-self-improving-school-system.pdf

Hargreaves, D. H. (2012). *A self-improving school system: Towards maturity*. Accessed online 13 June 2019 at: https://dera.ioe.ac.uk/15804/1/a-self-improving-school-system-towards-maturity.pdf

Harman, G. (1998). The management of quality assurance: A review of international practice. *Higher Education Quarterly, 52*(4), 345–364.

Hofman, R. H., Dijkstra, N. J., & Hofman, W. H. A. (2009). School self-evaluation and student achievement. *School Effectiveness and School Improvement, 20*(1), 47–68.

Janssens, F. J., & van Amelsvoort, G. H. (2008). School self-evaluations and school inspections in Europe: An exploratory study. *Studies in Educational Evaluation, 34*(1), 15–23.

Kyriakides, L., & Campbell, R. J. (2004). School self-evaluation and school improvement: A critique of values and procedures. *Studies in Educational Evaluation, 30*, 23–36.

MacBeath, J. (2005). Self evaluation: Background, principles and key learning. *Nottingham: NCSL*. Accessed online 3 June 2019 at: https://dera.ioe.ac.uk/5951/3/self-evaluation-background-principles-and-key-learning.pdf

MacBeath, J. (2006). *School inspection and self-evaluation: Working with the new relationship*. London: Routledge.

MacBeath, J., Boyd, B., Rand, J., & Bell, S. (1995). *Schools speak for themselves*. Accessed online 13 June 2019 at: https://www.teachers.org.uk/files/active/0/schools.pdf

Mahoney, M. J. (1977). Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research, 1*(2), 161–175.

Martin, K., & Jeffes, J. (2011). *Safeguarding children peer review programme: Learning and recommendations*. Slough: NFER.

Matthews, A. P., & Headon, M. (2015). *Multiple gains: An independent evaluation of challenge partners' peer reviews of schools*. London: Institute of Education Press.

NAHT. (2018). Improving accountability. *Report of the NAHT accountability commission. National Association of Head Teachers*. Accessed online 7 June 2019 at: https://www.naht.org.uk/news-and-opinion/news/structures-inspection-and-accountability-news/improving-school-accountability/

Nelson, R., & Ehren, M. (2014). *Review and synthesis of evidence on the (mechanisms of) impact of school inspections*. University of Twente.

Nelson, R., Ehren, M., & Godfrey, D. (2015). *Literature review on internal evaluation*. London: Institute of Education.

OECD. (2013). School evaluation: From compliancy to quality. In *Synergies for better learning: An international perspective on evaluation and assessment* (pp. 383–485). Paris: OECD Publishing. https://doi.org/10.1787/9789264190658-10-en.

Ryan, K. E., Gandha, T., & Ahn, J. (2013). *School self-evaluation and inspection for improving U.S. Schools?* Boulder: National Education Policy Center. Accessed online 2 June 2019 at: http://nepc.colorado.edu/publication/school-self-evaluation

Schleicher, A. (2018). *World class: How to build a 21st-century school system*. Paris: OECD. Accessed online 11 June 2019 at: http://www.oecd.org/education/world-class-9789264300002-en.htm

Shewbridge, C., Hulshof, M., Nusche D., & Stoll, L. (2011). *OECD reviews of evaluation and assessment in education: School evaluation in the Flemish Community of Belgium*. Paris: OECD Publishing. www.oecd.org/edu/evaluationpolicy

Stinton, J. (2007). *Evaluation of the college/school peer review pilot project* (Report no. RS7007). Norwich: The Research Centre, CCN.

Wenneras, C., & Wold, A. (2001). Nepotism and sexism in peer-review. In *Women, science and technology: A reader in feminist science studies* (pp. 46–52). New York: Routledge.

Dr. David Godfrey is an Associate Professor in Education, Leadership and Management at UCL Institute of Education in London and the programme leader for the MA Educational Leadership. He was co-director of the Centre for Educational Evaluation and Accountability until 2018 and was a lead inspector for the Independent Schools Inspectorate. An advocate of research-informed practice in education, his projects and publications include research-engaged schools, school peer review, inspection systems and lesson study. In July 2017, David was acknowledged in the Oxford Review of Education as one of the best new educational researchers in the UK.

Part II

Supplementing the Regional or National Accountability System

Chapter 2
Peer Reviews as a Complement to System Reviews in Queensland

Chris Diamond and Anetta Kowalkiewicz

Abstract In the Australian state of Queensland, voluntary peer reviews have evolved alongside the system's mandatory school reviews. Peer reviews involve the elements of self-diagnosis and collaborative reflection between schools, and have been supported but not controlled by the Department of Education's School Improvement Unit. In this chapter, new insights into the phenomenon of peer reviews are discussed. They are based on data from four cases studies of small groups of schools undertaking reciprocal peer reviews, and interviews with nine school principals and one principal supervisor. The study found that peer reviews provided a variety of benefits at the individual leader, school and system levels. These reviews are indicative of the professional maturing of the Queensland system which is embracing the benefits of collaboration, capability building and deprivatisation of school practice in order to drive and sustain school improvement.

2.1 Introduction

School accountability has been a consistent theme in educational research and practice over the last few decades. Education systems in different countries have different degrees of 'accountability pressure' (Altrichter and Kemethofer 2015, p. 37) which bring mixed results, including a number of unintended consequences associated with high pressure (ibid.).

In Australia, educational agreements signed by the federal, state and territory governments since 1999 have become more prescriptive, using key performance measures to monitor progress towards the achievement of national educational goals. Over subsequent years, however, the focus has shifted from strict accountability to an appreciation of school improvement which emphasises that improved student outcomes depend on improved school practices. This approach

---

C. Diamond (✉) · A. Kowalkiewicz
School Improvement Unit, Queensland Department of Education, Brisbane, Australia
e-mail: christine.diamond@qed.qld.gov.au

© Crown 2020

D. Godfrey (ed.), *School Peer Review for Educational Improvement and Accountability*, Accountability and Educational Improvement,
https://doi.org/10.1007/978-3-030-48130-8_2

acknowledges the complexity of the journeys towards the achievement of better student outcomes at individual, school and system levels. School improvement has been conceptualised by Masters (2016, p. 2) as

> … a continuous improvement cycle—a rigorous methodology for reviewing current school practices and outcomes; setting goals for improvement; designing and implementing school improvement strategies; monitoring changes in student outcomes; and reviewing and reflecting on the effectiveness of the school's improvement efforts.

In response to this shift towards school improvement, and as in many other countries, Australian states and territories sought to establish school improvement frameworks. The *National School Improvement Tool* (NSIT) (Australian Council for Educational Research 2012) and similar frameworks have been applied in external school review processes, and have provided policy levers to support system-wide improvement. However, the challenge is to sustain this change over time by 'building capacity within schools' (Centre for Education Statistics and Evaluation 2014, p. 3), as '[e]nsuring capacity for lasting improvement is critical to address challenges of quality and equality' (Stoll 2009, p. 116).

The critical question now is how this 'lasting improvement' in Australian schools can be achieved and supported by government. A 2018 think-tank report on the Australian government's role in education (Sonnemann and Goss 2018) found that neither a top-down, nor a bottom-up approach is sufficient for sustained improvement. Rather, '[a] more "adaptive" system design is needed, with stronger feedback loops for more systematic learning' (ibid., p. 9), as well as 'strong evaluative structures … to help embed evidence in practice' (ibid., p. 9).

In the Australian state of Queensland, the state government plays a vital role in supporting school improvement and developing schools' capacity to improve. In 2015, the Queensland Department of Education implemented a program of four-yearly school reviews (referred to in this chapter as system reviews). These reviews, conducted by the School Improvement Unit (SIU) located within the department but separate from the school delivery arm, aim to inform school strategic direction, operations and day-to-day classroom practice. However, in contrast to high-stakes school inspections conducted in the English system, system reviews provide formative feedback and support tailored to a school's context, without ratings against standards or sanctions. While mandatory and rigorous, Queensland system reviews do not feature elements that score high on established 'pressure scales' (Altrichter and Kemethofer 2015, p. 38).

Following the advent of system reviews, some schools sought feedback and reassurance that they were on their intended improvement trajectory in the period between these reviews. A variety of models of peer review were developed in a number of schools to evaluate the progress of school improvement initiatives. Some of these schools made requests to the SIU, which resulted in general offerings of process guidance and a reviewer brokerage service. This organic development of complementary reviews also attracted the attention of researchers within the SIU.

This chapter explores the phenomenon of voluntary peer reviews within the context of the Queensland state education system. Using data from interviews with nine principals and one principal supervisor of schools involved in peer reviews, as well as observations of four peer review processes, it aims to describe their scale, scope, characteristics and outcomes. In the absence of a universal definition, the term ‘peer reviews’ is used in this chapter to describe voluntary processes, initiated and driven by schools, which include elements of self-diagnosis and collaborative reflection between schools, with the consideration of system data, along with primary data collected by reviewers at the school site. The key themes identified from data are discussed, including the contribution to the contemporary debate about the potential for peer reviews to co-exist with and complement system reviews.

2.2 Peer Reviews and Internal Evaluation

The peer review phenomenon discussed in this chapter is, by its nature, a form of internal evaluation (or self-evaluation) that the literature generally defines as ‘something that schools do themselves, by themselves and for themselves’ (Swaffield and MacBeath 2005, p. 239). School self-evaluative practices have also been termed in the literature as collective review, self-review or self-assessment (Chapman and Sammons 2013). Kyriakides and Campbell (2004, pp. 27–8) identify four issues as important for the implementation of school self-evaluation:

- clarifying and developing consensus about the aims of evaluation
- creating an atmosphere of trust, openness and collaboration
- establishing protocols and procedures for the use of data
- creating criteria to measure school effectiveness (balance between broad and specific).

Internal evaluation, as opposed to its external counterpart, is usually conducted by teachers, other school staff, the principal, or by a special staff member designated by the school to serve as a school evaluator (Nevo 2001, pp. 95–6). Internal, self-directed evaluation has become more popular in response to the increased autonomy of schools and the distribution of responsibility for educational quality across various partners (Vanhoof and Van Petegem 2007, pp. 101–2).

External evaluation, on the other hand, has traditionally been linked to the notion of accountability, performed by the school district, the state department, a ministry or an independent evaluation consultant, and used to ensure schools are fulfilling their duties (Nevo 2001). Faubert (2009) provides a useful overview that differentiates between school evaluation schemes operating in Organisation for Economic Co-operation and Development (OECD) countries.

Both external and internal (or self) evaluations can be beneficial and have their limitations. Among the many benefits of internal evaluation are its empowering effects for (Nevo 2001):

- schools as organisations that can develop a self-monitoring ability and confidence in their educational direction

- teachers who collaborate and get involved in decision making outside of the classroom
- individuals who can acquire evaluation skills that are applicable in other contexts.

The major criticism of self-evaluation is the limited credibility of its findings (Nevo 2001, p. 97). Self-evaluation is seen as a necessary but insufficient ingredient to stimulate school improvement (Chapman and Sammons 2013, p. 2), but this shortcoming can be addressed by complementary external evaluation. Along with the growth of self-evaluation, the role of critical friends has been emphasised as 'helping schools see themselves from different perspectives' and 'broadening and deepening a school's self-knowledge' (Swaffield and MacBeath 2005, p. 239).

The potential co-existence of internal and external evaluations has been discussed widely in the literature. MacBeath (2008, p. 397) points out that self-initiated approaches to evaluation, '… driven by a natural desire for evidence' and improving practice, '… often had a vitality and drive because they were "owned" by schools themselves but also suffered because there was no… systemic support …'. Some schools use internal evaluation to prepare for external evaluation. External evaluation can provide a 'bigger picture' for the school, by using national standards and benchmarks. As Nevo (2001, p. 98) states '[e]xternal evaluation can stimulate internal evaluation, it can expand its scope and legitimize its validity' and internal evaluation can enhance external evaluation by providing a deeper local perspective.

Three models of how external evaluation and self-evaluation might co-exist in educational systems have been identified (Alvik 1996, cited in Swaffield and MacBeath 2005, p. 240, emphasis in original):

- *Parallel*: … two systems run side by side each with their own criteria and protocols.
- *Sequential*: … external bodies follow on from a school's own evaluation and use that as the focus …
- *Cooperative*: … external agencies cooperate with schools to develop a common approach to evaluation.

These models do not seem to exhaust the potential range of approaches. For example, another version of a sequential model, where the findings of external evaluation are used by schools as the focus of their self-evaluation, is also possible. These models of co-existence are also not mutually exclusive, and a combination of the three seems to be a viable alternative. An external agency can cooperate with schools to develop a common approach to evaluation that serves both the school's and the agency's needs (cooperative model), which will then be used in either the parallel or sequential model. While the distinction between parallel and sequential models refers to the implementation phase and is based on the process timeline, the cooperative model refers to the design phase and is about the way in which the evaluation process is developed before it is implemented. The discussion of the findings in this chapter will consider which model, or combination, the Queensland system and peer reviews represent and how they might co-exist so that the benefits could be achieved at both school and system levels.

2.3 Queensland Context

To support continuous improvement in the more than 1260 Queensland state schools and education centres, the state government, which regulates and operates schools, established the SIU and its program of school reviews (or system reviews) in 2015. School reviews are a new stage in the school improvement journey of the Queensland education system, as they view school improvement through the lens of collective responsibility, focused on providing schools with support, rather than a rating.

2.3.1 School Reviews

Using the NSIT as its guide to making judgements, the reviews identify areas for improvement and additional system support that may be required to improve student outcomes. As Fig. 2.1 indicates, the reviews provide substantiation for a school's new four-year strategic plan, which is '… an overview of a school's explicit improvement agenda … and clearly connects school priorities with the strategic direction of the [Education] department' (Department of Education 2018).

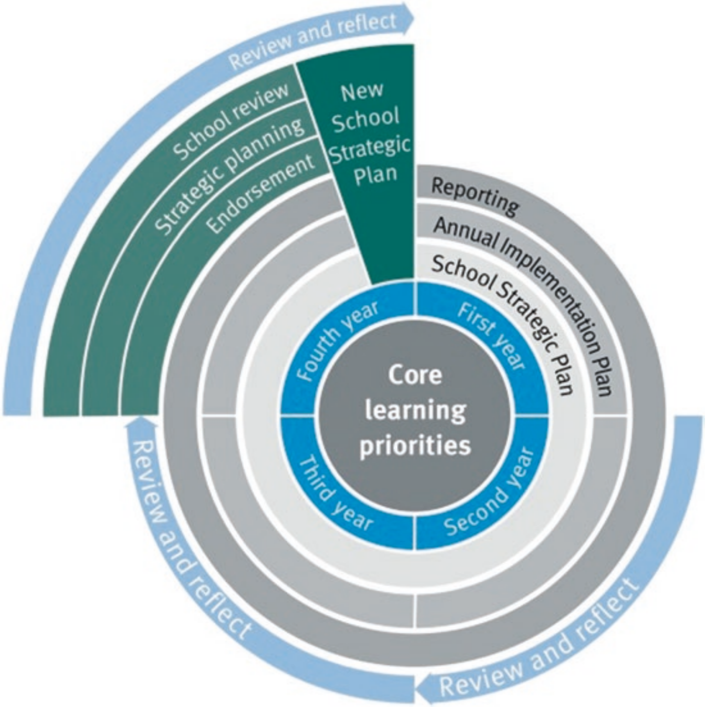

Review and reflect
School review
Strategic planning
Endorsement
New School Strategic Plan
Reporting
Annual Implementation Plan
School Strategic Plan
Fourth year
First year
Core learning priorities
Third year
Second year
Review and reflect
Review and reflect

Fig. 2.1 School planning, reviewing and reporting cycle for Queensland state schools © State of Queensland (Department of Education)

Each year, the SIU undertakes a desktop audit of an established set of ‘headline indicator’ school performance data to inform meetings between senior leaders of the SIU, regionally based principal supervisors and the regional director. Following these discussions a plan for the year ahead is developed, with schools due for review designated by the SIU for one of three review types (SIU 2018):

- full school: for most schools (the standard review type)
- priority support: for schools that require extra support
- self-determined: for schools on a positive improvement journey which are therefore given the autonomy to organise their own review.

The full school and priority support review teams are chaired by an SIU reviewer, an experienced principal trained in the use of the NSIT and review methods and temporarily seconded to the SIU. All current state school principals can undertake principal NSIT training and may then participate in the review of another school as a principal peer reviewer.

Full school review teams comprise of an SIU reviewer, a contracted reviewer (selected from a pre-qualified panel) and a principal peer reviewer. The reviews are undertaken at the school, over one to 5 days depending on the school’s size and complexity. Reviewers consider system data (such as standardised test results, report card results, school community satisfaction, attendance and behaviour data) and undertake extensive field work at the school, including interviews with school leaders, staff, students, parents and community members about how students at the school are taught and supported. The review team co-creates the review report and presents findings and improvement strategies, organised against the nine domains of the NSIT (Australian Council for Educational Research 2012):

- An explicit improvement agenda
- Analysis and discussion of data
- A culture that promotes learning
- Targeted use of school resources
- An expert teaching team
- Systematic curriculum delivery
- Differentiated teaching and learning
- Effective pedagogical practices
- School–community partnerships.

Priority support reviews proceed in the same manner as full school reviews, however the team consists of two SIU reviewers and one contracted reviewer, and the SIU continues to monitor the school for 12 months after the review (by visiting the school at 3, 6, and 9-month check-in points). This is to ensure the school is receiving the necessary support and is addressing the recommendations from the review.

Schools allocated a self-determined review have the autonomy to arrange their own review, with the SIU providing some funding and advice upon request. The SIU is often asked to undertake these reviews or the school may use an accreditation certification body, such as the Council of International Schools, with which it has a pre-existing relationship.

The SIU's training for principals aims to develop their understanding of the review process and how the NSIT can be used to support school improvement. The opportunity to be a principal peer reviewer is valuable professional learning where they, as part of a review team, collaboratively analyse another school's context and teaching and learning practice. In order to enhance the school improvement knowledge and skills in the system, the SIU also offers training in the NSIT to middle leaders, such as deputy principals and heads of department, and other school staff.

2.3.2 Pilot Health Checks

Following requests for assistance from schools interested in self-assessment, the SIU developed some guidelines, which were piloted in 2017 in four schools. This process (originally termed 'mid-point health checks', later 'school self-assessments') aimed to help schools evaluate the progress of their improvement agendas and to identify areas for further improvement between system reviews.

Based on the pilots, the SIU (2018, p. 84) recommends that the process:

- is owned by the school
- uses the NSIT (all or some domains) to promote alignment with the school review process
- includes school self-reflection and evaluation, combined with external peer assessment (preferably by colleagues trained by the SIU and experienced in the review process)
- is undertaken mid-way through a school's four-year strategic planning cycle, over one to 2 days, depending on the school's characteristics.

While not imposing a format, the SIU suggests that the mid-point health check team proceed in a manner similar to system reviews, that is, they would examine the school's performance data, interview school staff, students, parents and other stakeholders, and compile their feedback in a short report which suggests improvement strategies.

The SIU offers support in recruiting members for the health check team from the pool of current and former SIU reviewers, peer reviewers, accredited contractor reviewers or trained associate school leaders. On request, the SIU also provides training and other advice or support to potential reviewers.

2.4 Case Studies

2.4.1 Research Design

Emerging peer review processes provided the rationale for the SIU research project aimed at exploring this phenomenon. A case study research design was deemed most appropriate as it allowed us to 'investigate a contemporary phenomenon within its real-life context', using 'multiple sources of evidence' (Yin 1989, p. 23). Case study as a research strategy '… focuses on a single organization, institution, event, decision, policy, or group (or possibly a multiple set)' (Baker 1999, p. 321). Each case involved a peer review group consisting of between two and six schools, usually represented by the school principal, but sometimes by middle leaders as well.

Four peer review groups were included in the empirical study, which allowed for variation and cross-case comparison. Due to the exploratory nature of this research, and because the population under investigation is hard to identify and estimate, participants were selected and recruited through snowball sampling. Data were collected between May and November 2018. During that time, schools involved in peer review groups, as well as principals invited for interview, were at different stages of their peer review journey.

2.4.2 Method and Sample

2.4.2.1 Interviews

The most substantial data were collected through semi-structured interviews. One or both of the researchers interviewed each respondent, making reference to an interview guide and asking them to describe: their career to date, their involvement with system and peer reviews, and the peer review process and its outcomes. Some respondents were involved in a peer review as a reviewer and had their own school peer reviewed, in which case they were asked to reflect on similarities and differences between these processes. Interviews took between 30 and 60 min, and were recorded and transcribed.

In total, nine principals and one principal supervisor were interviewed (see Table 2.1). All respondents operated within (and to some small degree, across) the four peer review groups in which members tended to review each other's schools.

Table 2.1 Peer review groups by respondents

Peer review group	Respondent (pseudonym)	Role	SIU training and reviewer experience	Own school subject to peer review
1	Tanya	Principal of school A	–	✓ observed
	Amanda	Principal of school B	✓	✓ observed
	Michael	Principal of school C	✓	–
2	Susan	Principal of school D	✓	✓
	Claire	Principal of school E	✓	✓ observed
	Tom	Principal of school F	✓	–
	Barbara	Supervisor of principals of schools D, E, F	✓	N/A
3	Anne	Principal of school G	✓	✓
	Bill	Principal of school H	✓	✓
4	Peter	Principal of school I	✓	✓ observed

As Table 2.1 indicates, of the ten respondents, seven principals had their own school subject to peer review and nine respondents had been trained in and conducted system reviews in the past.

As some principals were involved in peer reviews before the study was undertaken, data collected during interviews included their reflections on these early experiences, covering the pilot health checks in the case of peer review group 3.

2.4.2.2 Observations

Interview data were augmented by observations of four peer reviews across three of the peer review groups as well as the observation of two meetings of the principals involved in peer review group 1. Field notes were taken and written up. Existing artefacts, such as reports from peer reviews, were also examined.

Table 2.2 describes the four peer review groups and the schools within them, and indicates which peer reviews were able to be observed by the researchers.

Table 2.2 Peer review groups by schools

Peer review group	School code	School location/type	School enrolment (2018, approx.)	Researcher observation of peer review
1	A	City/college	1700	✓
	B	City/primary	1000	✓
	C	City/primary	675	–
	J	City/primary	900	–
	K	City/primary	200	–
	L	City/primary	1125	–
2	D	City/primary	425	–
	E	City/primary	875	✓
	F	City/primary	450	–
3	G	Provincial/primary	550	–
	H	Provincial/primary	600	–
4	I	City/primary	625	✓
	H	Provincial/primary	600	–
	C	City/primary	675	–
	M	City/primary	675	–

Definitions:
City: located in an educational region close to the capital city
Provincial: located in an educational region some distance from the capital city
College: Prep (foundation) to Year 12
Primary: Prep (foundation) to Year 6

2.4.3 Analysis

Transcripts from interviews and field notes were analysed using NVivo 10 software. The analysis was intended to generate a pattern of inter-related themes which '… at the minimum describes and organizes possible observations or at the maximum interprets aspects of the phenomenon' (Boyatzis 1998, p. vii). Themes were derived inductively from raw data. A process of topic and analytic coding was applied during analysis. Topic coding involved '… creating a category or recognizing one from earlier, reflecting on where it belongs among … growing ideas, and reflecting on the data [being referred to] and on how they fit with the other data coded there' (Morse and Richards 2002, p. 117). The purpose of analytic coding was to further '…. develop categories theoretically' by '… not just linking them to the data but also questioning the data about the new ideas developing in the new codes' (ibid., p. 119).

2.5 Peer Reviews in Queensland State Schools

Existing research and literature on internal or self-evaluation focuses mostly on the benefits and the outcomes of the process. Less attention is paid to other aspects of evaluation practice, some of which may be critical to their outcomes. For example, the motivation of principals to seek feedback from peers on their school improvement progress, or to offer their own expertise and feedback to their peer colleagues, can significantly affect how the peer review practices are designed and performed and what outcomes they bring. Understanding the 'why' of peer reviews, that is, the motivations and needs of individual school leaders and school communities, can also help external agencies to (re)shape their approach to external reviews.

This section discusses the major themes inductively generated from the data. The identified themes describe different procedural aspects of peer review that may potentially determine the effects of the peer review on the school and those involved. They are grouped under two headings, rationale for peer review and process of peer review.

2.5.1 Rationale for Peer Review

The peer reviews subject to the study originated in different ways. Within the largest group, peer review group 1, which consisted of six schools, peer reviews evolved in response to a middle leadership training program delivered by the SIU. The group sought to arrange what they called 'snapshot reviews' in each school, where middle leaders could test their newly acquired skills. These peer reviews were seen as a 'win-win' by both middle leaders who could put their training into practice, and by the reviewed school that could get feedback on their school improvement progress. This initiative was also regarded as a way of 'revamping' the work of the geographically based cluster group that all schools belonged to, and which had stagnated over recent years.

Peer reviews within the other three groups were initiated by principals seeking feedback on school improvement progress between system reviews. Principals asked trusted peer(s) to review their school, which was later reciprocated. In peer review groups 3 and 4, principals also asked the SIU for an experienced reviewer to support the process.

Respondents identified a number of reasons for engaging in the practice of peer reviews. In all cases, the expectation was that peer reviews would identify the progress of school improvement between four-yearly system reviews, the ‘distance travelled’ since the last formal review. As Michael (principal of school C) explained:

> I think the schools that are keen to get feedback are schools that are, maybe, a minimum of 12 months, but maybe two years through their strategic plan following a review, so it’s that mid-point check-in process.

A few principals saw peer reviews as an opportunity to validate explicitly their perceptions of progress in the implementation of improvement strategies across the school. Principals emphasised the importance of ‘checking in’ with others, receiving ‘accurate feedback’ and ‘hearing through the viewpoints of others, their opinion and their findings’.

At the same time, some principals admitted they felt isolated in their role, which was not necessarily associated with a school’s geographic location, and how this detachment created the need for feedback. Tom (principal of school F) described principals and their need for feedback as follows:

> … principals are ‘isolates’. They’re sitting there trying to strategise everything in their heads but they also do need some input from people outside their school, and that has to be a volunteer thing and the more you get to do it, the more they’ll see how it helps them move forward.

Peer reviews were often seen by the respondent principals as complementing system reviews. This, as Tom (principal of school F) explained, was particularly important for schools that had a priority support review and the subsequent 12 months of support from the region and the SIU:

> … It is very intense in … priority school reviews, as you probably get some feedback, and we were very successful. But the trouble is when you come off one of those, where to then? That’s the big void that sits around these schools.

Bill (principal of school H) believed their peer review was critical in assisting the school to adjust their strategic priorities in response to a changing school context and other challenges in between system reviews:

> Students can change every term or six months … So you all of a sudden go ‘well our priorities have to change’ to be responsive to that shift in dynamic clientele … In four years’ time you could have 100 per cent staff changeover … That’s why I think those mid-reviews, those health checks … are a good strategy.

Some principals intended to use the peer feedback as leverage to implement specific change in their schools. They had learnt about the powerful effect of system reviews and sought additional support to pursue change. Claire (principal of school E) emphasised ‘the importance of having that message [rationale for change] coming from elsewhere, having a bit of an imperative around this', while Susan (principal of school D) was even more explicit, talking about 'using the [peer review] report as leverage to move reluctant people forward'.

A number of respondent principals anticipated that peer reviews, by enabling school staff to celebrate achievements, would enhance morale and re-engage staff in the improvement agenda.

In some instances, the key motivation for conducting peer reviews was developing the capability of staff, particularly middle leaders. This was the case for peer review group 1 which provided NSIT training before the peer review process, but the potential benefits of professional learning from peer review were also acknowledged across other groups. According to Susan (principal of school D), involvement in peer reviews allows participants to better understand that they are part of a 'bigger picture of school improvement … [t]here's nothing better than giving people the tools, magnifying glass or way to see that big world'.

Finally, for peer review group 1 in particular, the development of authentic relationships among school leaders, which would further facilitate 'working together in a trusting environment', was an important motivation.

2.5.2 Process of Peer Review

Decisions about the procedural aspects of peer reviews were the remit of the principal whose school was being reviewed. Although different aspects were often subject to group discussion, it was ultimately the principal who determined the peer review focus and arranged reviewers and review activities to be conducted in their school.

2.5.2.1 Pre-review Training

In three of four peer review groups, specific training was arranged for review participants. Alternatives included sending staff to SIU training, inviting SIU trainers to undertake bespoke sessions or having the training delivered by the principal supervisor (who had previously been an SIU trainer). In all cases, training focused on the use of the NSIT and the associated conceptual framework. In the case of peer review group 3, training was not needed as all participants were experienced SIU reviewers.

2.5.2.2 Review Scope

In general, the peer reviews focused on the areas or initiatives that the school had been working on following their last system review. In some cases, direct links and comparisons were made between new findings and those from that review. This narrow focus usually involved one to three domains of the NSIT (unlike system reviews, which encompass all nine domains).

The reviews conducted by those experienced in the system review process tended to be more closely aligned to the NSIT. Respondent principals articulated the benefits of applying the NSIT as the promotion of a common language and guide to making judgements, and in terms of setting a standard for, or pathway to, effective practice. As with system reviews, peer reviews did not rely on an established or standardised set of questions to be used by all reviewers.

In each school in peer review group 1, the focus was on one particular improvement initiative from their last system review report and the extent to which it had been consistently implemented in classrooms. In school E from peer review group 2, Claire sought to check implementation of strategies against the school's action plan (a step following a system priority support review), rather than whole domains. In school I, the progress was discussed for each of the school's three priorities in relation to each of the nine NSIT domains.

2.5.2.3 Reviewers

At the time of the study, two distinct methods were used to select peer reviewers. One was to identify those who had previous experience in conducting system reviews and associated expertise in the NSIT. This approach was taken by those who had knowledge of, and trust in, system reviews and wanted their peer review to be conducted in a similar way, clearly linked to NSIT domains. On request, some schools were provided with SIU reviewers. Bill (principal of school H) explained that the benefit of experienced reviewers was that they are 'external people who had some real knowledge of the school improvement tool [and] other schools, but no real connection with our school'. Other principals asked principals they knew and had been trained, or their supervisor.

The second approach was to select school staff (usually aspiring middle leaders) who had minimal or no training in the NSIT and no review experience. This method was used when one of the key purposes of the peer review was to build staff capability. In the case of peer review group 1, principals of participating schools chose one person who had expertise in the area of review focus (so that the reviewed school could benefit), and one person who could learn substantially from such an experience to participate as reviewers. This ensured that all schools involved in the peer review benefited.

A combination of the two approaches was also evident, with experienced reviewers being shadowed by aspiring leaders from participating schools. Some schools also included teachers as peer reviewers, which Claire (principal of school E) explained in terms of the 'power in teachers talking to teachers' to bring 'a sense of credibility around the messages'.

Participants in the peer review processes examined in this study also identified the need to determine the optimal size of review teams. Having what Tom (principal of school F) called 'external eyes' to look and reflect on the situation at a school was seen as advantage, but at the same time risked gathering too many interpretations, making it difficult to agree on the key findings and recommendations for the school to focus on in the next stage of their improvement journey. This is reflected in Tanya's (principal of school A) comment:

> Well, we might get to a point where we think we've got too many people going into schools, for example with a small school. It will be interesting to see how it evolves. It may not always be the group but until we get a confidence level, only then we might scale it back from being the large group to a small group.

2.5.2.4 Activities

Interviews and observations demonstrated that the review process was shaped according to the needs and preferences of the school being reviewed, and this was viewed as a benefit. In some cases, review activities were strongly influenced by the reviewers' experience with system reviews. As a result, each peer review was conducted in a different way, but the following common activities were apparent:

- analysis of documentation and data provided by the school (before the review or on the first day)
- interviews with school staff in accordance with a timetable provided by the school, often including a relief teacher to cover the interviewee's class
- discussion of findings by reviewers as a group and co-writing of the report and recommendations
- exit interview with the principal, with an opportunity for the principal to give feedback about the process and findings
- sharing the key findings with school staff.

Additionally, in some instances the leadership team presented to the review team their reflections on the recommendations from the most recent system review, and group interviews were conducted instead of individual conversations. During peer reviews in schools peer-review group 1, staff from the peer-reviewed school observed the final discussions of the reviewers leading to the review findings. This was seen as a valuable opportunity to 'hear the feedback unsanitised' (Amanda, principal of school B).

2.5.2.5 Report

In almost all peer reviews, a report was drafted after interviews and discussion. In comparison with system review reports which are up to 30 pages in length, peer review reports were shorter and more succinct. In one case, only a half-page summary was provided. The general structure was to provide positive affirmation of the successful work, and areas for improvement. For school G, the report also explicitly referred to the recommendations from the most recent system review and school self-assessment. Explicit links to NSIT domains were found in a few reports across the sample. In all cases, the report was produced promptly, often being provided at the end of the review.

In peer review group 1, by contrast, a collaborative decision was taken not to produce a report per se. Instead, commendations and recommendations were collated on a whiteboard during discussions and were typed and provided to the school principal.

In a few cases, reviewers were asked to write reports in ‘teacher-friendly language’ (Susan, principal of school D). This was important for Tom, principal of school F, who wanted ‘some accurate information that makes sense for teachers’.

2.5.2.6 Follow Up

In most cases, findings from the peer review were used to inform the next steps in school improvement. Leadership teams, with school staff, often scrutinised the reports and discussed what they would do in response. In Tom’s school F, findings were discussed in close comparison with the findings from the latest system review, which facilitated a process of regular monitoring of improvement progress. In Susan’s school D, at the end of every term, a group of curriculum leaders referred to both sets of recommendations when considering where they were at and then teachers used a similar process as part of self-reflection.

2.6 Benefits of Peer Reviews

A range of outcomes from peer reviews were noted by respondents. Peer reviews facilitated school improvement by highlighting progress and areas for additional work. They provided professional learning for peer reviewers and validated principals’ perceptions regarding the current school improvement and its future direction.

2.6.1 Facilitating School Improvement

Peer reviews were seen by principals as helping schools to identify progress, strategically manage school improvement and bridge the time between system reviews. It was emphasised by some respondents that the reviews were not a token activity, but a genuine process they trusted.

Respondents identified that, in comparison with system reviews, peer reviews provided more detailed feedback on the progress of improvement, what had already been achieved, and what inconsistencies were still evident in practice. As Bill (principal of school H) explained:

> … when you do the initial review it's around the diagnosis not the medication and you're just really focusing on 'here are some problems of practice that you might want to consider to move forward' … The mid-cycle review you're able to do a little bit of both. So we were able to go 'yeah, you're heading down that track but I can also see this stuff and maybe that's a way of moving forward' … It's not to say that the feedback they're giving is the right feedback, but it allows [us]… to probe and ask the question and provoke the curiosity to go to the next step. I think that's what that does, so that people don't spend that four years just trying to fit what they already do in.

All of the schools in the study that had a peer review used their findings to inform their strategic direction. Claire (principal of school E) called it creating 'a point of truth'. In school D, peer review findings informed planning and professional learning agendas for the following year. According to Barbara (principal supervisor), the benefits of the peer review included: 'a good space to check in, encouraging regular engagement with the strategic plan [and] permission to play with and adapt [the] strategic plan'.

In some instances, peer review highlighted the achievements in school improvement, which were then celebrated. This was seen as crucial to maintaining the momentum of the improvement journey. It was, according to Bill (principal of school H): 'that real shot in the arm of "we're doing a great job and let's go again"'.

According to respondents, peer review also helped school communities to prepare for their next system review. School staff were looking forward to the review as they were 'armed' with the information they had obtained through their peer review, and they wanted to demonstrate growth and make comparisons.

2.6.2 Professional Learning

There were elements of professional learning identified in all of the peer reviews in the sample. The process was seen as a great learning opportunity, mostly because of the reflection it prompted from its participants. As Amanda (principal of school B) said, 'if you're not driving home thinking how this might affect or help my own school and my own job, then it's a waste of time'.

School staff participating in peer reviews saw them as a great professional learning opportunity. Feedback from principals and middle leaders from schools involved in peer review group 1, and Anne (principal of school G), identified this opportunity as 'the best PD [professional development] they had ever done' and 'the best PD for principals ever that Queensland has offered, while still on the job'. Tanya (principal of school A) suggested that the peer review experience would be valuable professional development for aspiring leaders. It was also a common perception that school staff learnt a lot during the process of preparing for a peer review.

Respondent principals identified the following areas where their knowledge or skills had been improved:

- increased knowledge and understanding of the practices implemented in other schools, including those in their cluster (in peer review group 1, the idea of hav-

ing an additional session to share more details about some practices was considered)
- ability to diagnose the situation at a school, ‘being able to pick up what is going on in a school very quickly’ (Claire, principal of school E)
- leadership skills, particularly for middle and/or aspiring leaders
- review skills, such as questioning techniques, interview protocols and using the NSIT to generate findings
- ability to respectfully give and receive feedback
- data literacy through group discussions of school data.

2.6.3 Validation of Principals’ Perceptions and Plans

Many principals reported that the value of the peer review was the validation of their own perceptions about the school, the improvement process and its future direction. As Claire (principal of school E) explained, it was about ‘being able to … check in that my own assumptions were actually based in fact’. Claire also saw the benefit of independent evidence of school improvement progress:

> I know there’s always the question coming from central office around ‘how do you know?’ This was our evidence for ‘this is how we know’. We thought we knew but [now] we really know, which is really nice. So it sort of gave us that confirmation that the measures that we were using were actually valid and reliable.

In most cases, the peer review confirmed what principals said they already knew, and according to Bill (principal of school H), as a result of this validation, he discovered his school had improved to even a greater extent than he believed. In some schools, the peer review provided revelations or ‘a-ha moments’ and pre-empted issues, described by Amanda (principal of school B) as ‘little things that could potentially be big things’ if they were not identified and addressed.

2.7 Values Prevalent in Peer Reviews

While varied in their implementation, the peer reviews in the study were identified as based on and promoting a range of common values. These values—flexibility, authenticity, trust, credibility and rigour—are consistent with the philosophy of system reviews.

There was considerable flexibility in the peer review process, which meant that each review was tailored to the school and principal’s context and needs. They varied in duration, formality and focus. This flexibility was strongly related to the voluntary nature of peer reviews, that it was a professional choice, and the fact that they were undertaken by principals interested in continuous school improvement. Such flexibility aligns with the strong and shared belief of principals that they know their schools (students, teachers, practices) and therefore are best able to shape the review to the needs of the school. While such principal control could potentially lead to confirming the favourable perception of the school rather than challenging the school's status quo, other values—of authenticity, trust, credibility and rigour—are likely to counter such a possibility.

Peer review was viewed by those involved as an authentic and genuine process. Peer reviews provided opportunities to build sincere relationships between schools, as well as between the individuals in those schools and clusters. It was the authenticity of the peer review process that prompted many review team members to identify it as 'the best professional learning development' they had ever participated in. This finding is consistent with discourse in the literature, which indicates the need for authentic professional learning and 'reframing professional development: shifting from a focus on delivery of content, to support for professionals as they inquire into and adapt their practices in the contemporary workplace' (Webster-Wright 2010, p. 1).

The desire for genuine, professional conversations, feedback and learning, was evident across the peer review groups, although most significantly in peer review group 1, which included more inexperienced peer reviewers when compared to other groups. The review protocols agreed within this group emphasised the value of authentic professional development opportunities for middle leaders, of building authentic networks, and of schools sharing, learning and growing together. As Amanda (principal of school B) noted:

> … having conversations with other people and deciding on what they think about a situation, their feedback on a situation, there is so much real conversation happening.

The building of authentic, trusting relationships with other principals was also a specific goal and achievement for this group. Peer reviews enhanced relationships which were then further built upon, with staff visiting each other's schools and having further conversations regarding school improvement.

The commitment to authenticity meant that respondent peer reviewers sometimes went beyond simply identifying issues, to suggesting solutions based on their experience as principals and as peer reviewers.

Trust was a key element of peer reviews. Some principals involved in the peer reviews had worked together before, or they had been part of formal or informal networks, mentorships or learning communities, or formal SIU review teams. Such pre-existing relationships meant they had already built trust, which provided a basis for peer review work. The presence or absence of trust often determined the principal's engagement in a peer review process. As Claire (principal of school E) reflected:

> Trust is a massive thing, inviting another principal into your school. I guess it's different when you have a [system] review because the people who come in around a proper formal review you know they've all been trained, you know they come under a particular mind-set and you're guided by the people from the SIU who have very strong parameters around how they work. … To do it with your peer, you've got to know you're not going to get thrown under the bus the next time they have a conversation with someone.

There needed to be trust in shared intent: as Peter (principal of school I) said, 'the school reviews are really good because they're purposeful'. Principal respondents were particularly mindful that the purpose and 'essence' of the review process be respected and not 'watered down'.

While principals had trust in their peers and the process to begin with, the peer review process helped to enhance trust among its participants. Having previously gained trust in the SIU review process and its effectiveness, the principals tended to have more trust in the peer review process the more similar it was to the system review process. Trust in reviewers and their expertise was also critical: while many principals seemed to have higher trust with the more experienced, SIU-trained reviewers, some pointed out that for their teaching staff, it was often important that they were interviewed by peers (that is, other teachers, who were included in some peer review teams). Finally, it was essential that school staff trusted the peer reviewers so that they would open up and share their practices and perceptions.

The credibility of the peer review was enhanced when there was strong reviewers' expertise and experience, as well as explicit reference to the NSIT (due to confidence in the tool). The potential for unreliable or invalid findings was reduced by involving reviewers with SIU experience or those known from existing professional networks, who were from outside the school and, on occasion, by increasing the size of the review teams (although this approach was not supported by all respondent principals). Anne (principal of school G) strongly advocated for a high level of experience in peer reviewers:

> I think that experience level shouldn't be undervalued and therefore privileged so that would be my preference. So if I had a mid-cycle review again I would check with the [SIU] and see if we can get someone who can come in just to be sure.

The standing of peer review was also affected by how its findings were generated and reported back to the school. A range of approaches were evident and depended on the reviewed school's needs and expectations. More formal approaches involving structured and systematic analysis and reporting, as well as explicit links to the NSIT domains and recommendations from previous system reviews, provided more detailed feedback and strategic guidance for schools. However, in some peer review groups (specifically peer review group 1), the professional development aspects of the peer review appeared to be as important as the feedback to the school, which as a result had been generated with less rigour.

2.8 Peer Reviews and System Reviews

It is evident from interview and observation data that the peer reviews have been inspired by the SIU's system reviews, and there is a high level of commitment among respondents to the system's broader school improvement agenda. Our study revealed many links between the peer reviews and system reviews, although peer reviewers emphasised that the intention was not to make the peer process look like the system reviews, but instead to apply those elements that were seen to enhance the review's rigour. The elements of system reviews that were promoted and adopted in peer reviews (to varying degrees) included:

- analysis of school documentation and data before and at the start of the review
- presentation by school staff at the start of the review
- an appreciative inquiry approach
- data collection via non-standardised interviews with a range of school stakeholders
- review team discussions between data collection opportunities
- reference to the domains of the NSIT.

Principals' reliance on the NSIT as a useful, evidence-based and widely tested tool was an integral part of the frame of reference of many peer reviewers. Even in those cases where the links to the NSIT were not explicit, its conceptual framework guided conversations about school progress among peer review teams. It also permeated collective sense-making and provided a common language during peer review interviews and discussions. The principal of school F, Tom, commented:

> It's a reference we can use, it's something we know about, we've been through it. Otherwise it's going to leave us open to interpretation when the real review happens.

References to system reviews were frequently made by respondents during interviews and were common throughout the four observed peer review processes. The peer reviews were driven by individuals whose thinking was shaped to a great extent by SIU training, the NSIT, and their experience as reviewers in system reviews. Such associations are reflected in Susan's (principal of school D) comment:

> … the majority of principals I speak to who have done the training and then gone and done a few reviews, they are the ones that are the biggest advocates and will turn to each other and ask each other to submit feedback.

Due to these links between them, peer reviews and system reviews represent two potentially compatible processes that could more than co-exist—they could enhance each other and together contribute to the continuous school improvement in Queensland state schools. The principals interviewed saw this potential. For example, Amanda (principal of school B) explained:

> I'd like to think that maybe in a couple of years' time that the [system] review team says 'these cluster schools that are doing this stuff are actually using the tool, and they don't drag it out once every four years and go what do we need to do, it's actually ingrained in their language and they're revisiting it every term'.

The review tool plays a significant role in integrating peer and system reviews. Another element that brings the two types of reviews together is the reviewers. The principal of school F, Tom, noted the need to create a database of trained and experienced reviewers who could be matched to a school's context and needs:

> So when you have peer reviews, you could have a list of people, you might see some fantastic principals who work with [the SIU] and you say 'I would recommend these people if you want to do some peer stuff, they're great peer people and they'll work with you and

> your teams for whatever context you want'. I think that is better. At the moment I have to search my own database of what I know, and there are only two names I have and if they aren't available …

Such a database would support peer review practice and it could potentially be created by the SIU, which has trained and worked with hundreds of reviewers.

Finally, another principal suggested peer reviews would benefit from some level of moderation being applied to the process. Anne's (principal of school G) comment suggests such support from the SIU would be valued:

> … if the SIU formalises it a little bit more then I think this idea of mid-cycle reviews, phone a friend … then I think more schools would go down this track. I think there's something to be said about the SIU still having a role in … these mid-cycle reviews, just as moderation.

Such quality assurance could make the process and its outcomes more consistent and comparable between schools and across the system.

The continued legitimacy of the SIU's role, and the complementary relationship between the system and peer reviews, was underscored by Tanya (principal of school A):

> I think we will still need a central body, for consistency across clusters. Ours would look different to others. There will still be a need for centrally located reviewers to come in and speak to us about how we are in relation to the state.

A key question for the future of peer reviews in Queensland is whether both systems continue to run in parallel, and cooperate to apply the common review tool in the same way in order to achieve consistency and, over time, constitute a sequential model where system reviews are being informed by the results of peer reviews. It seems that Alvik's (1996, cited in Swaffield and MacBeath 2005) three ways of co-existence (parallel, sequential and cooperative) could be blended. For the SIU, this would mean greater comparability of results across the schools, should peer reviews inform the system reviews in the future. For individual peer review schools, this would increase the rigour and credibility of the process, although potentially compromising its authenticity.

2.9 Conclusion

Overall, peer reviews offered a variety of benefits at the system, school and individual levels. At the school level, peer reviews informed the strategic direction of school improvement. The feedback received between system reviews about schools' current state (both achievements and shortcomings) and next steps, assisted schools to adjust their improvement approach to better use existing opportunities. Peer reviews also promoted school self-reflection. They encouraged school leaders and staff to have collaborative conversations using the common language of school improvement provided by the NSIT.

At an individual level, peer reviewers (principals and middle leaders) appreciated the authentic professional learning opportunity these reviews provided. Principals of schools that were peer reviewed had additional opportunities to enhance their skills in self-assessment, as well as giving and receiving feedback. All of the principals valued the professional networks that grew and deepened as a result of peer review, which could be tapped into for peer review and other purposes.

Peer reviews fostered and highlighted the professional development of school principals. A prime motivation for the organic growth of peer reviews in Queensland was that individual principals had sought to collaborate with their peers in relation to school improvement rather than acting alone. Principals' self-perceptions (perhaps affected by feelings of professional isolation) have contributed to this desire to collaborate through peer review. Hence, peer reviews more broadly reflect principals' maturation as school leaders. As Tom (principal of school F) said, '… it's a natural progression of principals who want the improvement and recognise you can't do it by yourself'.

In addition to the benefits to principals as leaders and education professionals, and to school communities, peer reviews also benefited the Queensland state school system. Principal supervisor Barbara, who was an advocate for and participant in peer reviews, was able to expand her knowledge of the schools she supervised and their leadership teams. This also allowed her to identify needs that could be met by regional resources. Peer reviews also provided an opportunity to work with other principals she supervised and get to know their ways of working.

For the SIU, peer reviews reflected the quality and impact of its work, the efficacy of training, the value of system reviews and the effectiveness of communication related to system reviews. The peer reviews also 'smoothed the way' for system reviews—school leaders and staff were familiar with the process (and for many, personally and professionally committed to the inquiry way of working) and were able to pre-empt what the system review might find and recommend. These reviews helped to maintain the momentum between system reviews. As well as reassuring the schools that they were on the right path, peer reviews enhanced the capability of principals to lead sustained school improvement.

Developing a shared way of applying the NSIT in review and a common pool of potential peer reviewers could promote greater consistency and comparability of peer reviews. The application of the tool would need to be determined in collaboration with school communities, and be flexible enough to suit various schools' contexts and needs. Using such an approach, applied by trained and experienced peer reviewers, peer reviews could continue to be performed independently of system reviews, providing valuable findings to inform the system process on-demand.

In the future, peer reviews and system reviews could be part of a consistent model, where both forms will build upon each other's work. Some principals involved in peer reviews have already declared they intend to present peer review findings to system reviewers as evidence of their school improvement progress and self-diagnosis. It is anticipated that the SIU will continue to support peer reviews but that principals, as system leaders, will be central to promoting an approach that benefits from both consistency and flexibility.

References

Altrichter, H., & Kemethofer, D. (2015). Does accountability pressure through school inspections promote school improvement? *School Effectiveness and School Improvement, 26*(1), 32–56.

Australian Council for Educational Research [ACER]. (2012). *National school improvement tool*. State of Queensland (Department of Education, Training and Employment) and the Australian Council for Educational Research.

Baker, T. (1999). *Doing social research*. McGraw-Hill College.

Boyatzis, R. E. (1998). *Transforming qualitative information. Thematic analysis and code development*. Thousand Oaks/London/New Delhi: Sage.

Centre for Education Statistics and Evaluation [CESE]. (2014). *School improvement frameworks: The evidence base*. Sydney: NSW Department of Education and Communities.

Chapman, C., & Sammons, P. (2013). *School self-evaluation for school improvement: What works and why?* Berkshire: CfBT Education Trust.

Department of Education. (2018). *School planning, reviewing and reporting framework: The improvement and accountability framework for Queensland state schools*.

Faubert, V. (2009). *School evaluation: Current practices in OECD countries and a literature review* (Vol. 42, OECD Education Working Papers). OECD Publishing.

Kyriakides, L., & Campbell, R. (2004). School self-evaluation and school improvement: A critique of values and procedures. *Studies in Educational Evaluation, 30*(1), 23–36.

Macbeath, J. (2008). Leading learning in the self-evaluating school. *School Leadership and Management, 28*(4), 385–399.

Masters, G. (2016). *Schools as learning organisations*. Australian Council for Educational Research.

Morse, J. M., & Richards, L. (2002). *Readme first for a user's guide to qualitative methods*. Thousand Oaks/London/New Delhi: Sage.

Nevo, D. (2001). School evaluation: Internal or external? *Studies in Educational Evaluation, 27*, 95–106.

School Improvement Unit [SIU]. (2018). *Improving student outcomes through school reviews: A toolkit for principals*. School Improvement Unit.

Sonnemann, J., & Goss, P. (2018). *The Commonwealth's role in improving schools*. Grattan Institute.

Stoll, L. (2009). Capacity building for school improvement or creating capacity for learning? A changing landscape. *Journal of Educational Change, 10*, 115–127.

Swaffield, S., & MacBeath, J. (2005). School self-evaluation and the role of a critical friend. *Cambridge Journal of Education, 35*(2), 239–252.

Vanhoof, J., & Van Petegem, P. (2007). Matching internal and external evaluation in an era of accountability and school development: Lessons from a Flemish perspective. *Studies in Educational Evaluation, 33*, 101–119.

Webster-Wright, A. (2010). *Authentic professional learning: Making a difference through learning at work*. Dordrecht/Heidelberg/London/New York: Springer.

Yin, R. K. (1989). *Case study research: Design and methods*. Newbury Park: Sage.

Dr Chris Diamond is the Manager, Policy and Research, School Improvement Unit, Department of Education, Queensland, Australia, where she leads research into school improvement. She joined the department in 2010 and has worked in various strategic policy and research roles. Earlier in her career, Chris held research roles within the Academic Board Office and the (then) Department of Government at the University of Queensland. In 2003, she was awarded her PhD in political science at the University of Queensland, receiving a Dean's Commendation for Outstanding Research Higher Degree Theses. Chris then took up an Australian Research Council Linkage postdoctoral fellowship in the School of Political Science and International Studies. She has published widely in the areas of gender, work and work/family balance.

Anetta Kowalkiewicz is the Principal Research Officer in the School Improvement Unit, Department of Education, Queensland, Australia, where she is undertaking research in the area of school improvement and educational leadership. Before this, she taught at the University of Queensland, the Queensland University of Technology and Griffith University in Brisbane. She was awarded a PhD in organisational studies from the University of Economics in Poznan, Poland, and was an Australian Research Council Linkage postdoctoral fellow at the University of Queensland Business School, where she investigated organisational change and culture change in a large public service organisation. Her current research supports the department's strategic direction and informs further improvement in Queensland state schools.

Chapter 3
'Layering' Peer Enquiry as a System Change Strategy: Some Lessons from Wales

Mark Hadfield and Mel Ainscow

Abstract This chapter describes a peer enquiry process developed by headteachers in Wales. Peer enquiry was one of the key collaborative school improvement approaches used within a regional challenge that aimed to develop a more 'self-improving system. The chapter discusses the potential of peer enquiry for bringing about system change by gradually replacing 'vertical' accountability structures by more collaborative and reciprocal 'horizontal' approaches. The key differences for those involved in undertaking a peer enquiry, rather than a peer review, arose as much from the context in which it took place as they did from any substantive differences in the processes used.

3.1 Introduction

Like many successful school systems, Wales is small, with only 1569 schools.[1] Despite its size, however, the performance of the school system is a cause for concern, particularly in terms of outcomes for learners from low-income families. Although it is part of the United Kingdom, Wales has a form of self-government, the National Assembly, created in 1998 following a referendum. There are 22 local authorities responsible for a range of public services, including education. A key turning point in the development of the system was in 2009, when Wales was ranked the lowest of all the countries in the PISA rankings of the UK. This resulted in a raft of reforms to the education system, including the merging of the school

---

[1] https://www.besa.org.uk/key-uk-education-statistics/

M. Hadfield (✉)
Urban Programmes Research Group, Colyton, UK
e-mail: appliedethno@btinternet.com

M. Ainscow
University of Manchester, Manchester, UK

© Springer Nature Switzerland AG 2020

D. Godfrey (ed.), *School Peer Review for Educational Improvement and Accountability*, Accountability and Educational Improvement,
https://doi.org/10.1007/978-3-030-48130-8_3

improvement services of the 22 local authorities into four regional consortia. Representatives of the consortia worked with central government on the design of a National Model for Regional Working. This was seen as a radical change, in that it envisaged a more collaborative partnership-based approach.

A country review of the education system conducted by OECD, carried out in 2014, (OECD 2014), argued that whilst the pace of recent reform had been high, it lacked a long-term vision, an adequate school improvement infrastructure and a clear implementation strategy that all stakeholders shared. Most significantly, the review argued that national policy still placed too much emphasis on accountability and not enough on providing support. In response the Welsh government published a five-year reform plan (DfES 2014). A key element of the plan was the Welsh government's interpretation of a 'self-improving school system', which involved:

- Transforming school improvement from being something that was once 'done to' schools to something that is being 'done by' schools.
- An end to the top-down improvement 'service' being delivered to schools and instead seek to empower school leaders to work together, taking control of their futures and their development.
- Those within schools taking responsibility for raising standards within their own organisations.
- A strengthening of the partnerships between schools, such that they are able to support and challenge one another. (DfES 2014 p. 21)

At the heart of Wales's approach to accountability is its' National School Categorisation System. Annually schools in Wales are evaluated by Challenge Advisors (CA) from the regional consortia and placed into one of four support categories: green, yellow, amber or red. Wales has its own national inspection service, Estyn which operates on a 7 year inspection cycle. Estyn inspections cover a broader range of areas than 'categorisation'.

This chapter describes and analyses the development of a peer enquiry process that was a key element of one of the regional consortia's improvement strategy, known as the Central South Wales Challenge. In so doing it throws light on the problems that occur when trying to introduce a strategy that necessitates collaboration between schools, as part of the move towards a self-improving system, within a policy context constrained by high stakes accountability structures, such as 'categorisation'.

3.2 The Central South Wales Challenge

The Challenge was developed by the Central South Consortium (CSC), one of the four regional consortia in Wales that had taken over responsibility for school improvement and professional learning provision from local authorities in 2014. It involves a partnership of just over 400 schools, across five local authorities, including the capital city, Cardiff. It was instigated by the directors of education and received the endorsement of local politicians, providing the mandate that seems to have been so important to the success of previous regional UK Challenges in London and Manchester (Ainscow et al. 2019). The Challenge involved four linked strands of collaborative activities, including peer enquiry. Each strand, as discussed later, varied in the intensity of the collaborative work they entailed and so were open to a range of schools. Each also aimed to replace or adapt aspects of the previous local school improvement infrastructure in order to develop a more ‘self-improving’ system.

The Challenge had to work within existing national accountability structures in which schools are at the ‘bottom’ of a three-tier system and at the top of which sits central government. In the middle are the four regional consortia, working in partnership with local authorities. The Challenge faced the issue of how to create a more collaborative regional approach to school improvement, whilst working within a national accountability structure premised upon comparative judgments of school performance that generated a degree of competition between school leaders. Each of the four strands focused on a different aspect of the existing system and each adopted a different approach to system change.

The peer enquiry process, as it developed over a four-year period, faced a particular set of challenges. Unlike other strands of the Challenge, where the consortium could gradually replace existing forms of provision and structures with more collaborative structures, peer enquiry had to be ‘layered’ (Mahoney and Thelen 2011) over of an existing national accountability system. Layering, in this instance, meant promoting a new, and more, collaborative approach in the hope that if it persisted, and spread, it would gradually reduce the political and cultural significance of the existing hierarchical structures. Peer enquiry was an attempt to normalise more ‘horizontal’ (Suggett 2014) approaches to accountability, based upon judgments by professional peers and other stakeholders in the local education system, rather then ‘vertical’ approaches, based upon ‘independent’ inspection services funded by central government. The development of ‘horizontal’ approaches to accountability within a more collaborative system was part of a broader attempt to fundamentally change its cultural dynamics. A dynamic based around lessening the cultural significance of external judgments, such as a schools’ ‘categorisation’ or outcomes from a periodic Etsyn inspection, on leaders professional status and esteem, whilst increasing the value attached to their engagement in collaborative school improvement processes. The story of the development of the peer enquiry process in CSC was therefore in part the role played by more ‘horizontal’ approaches in transitioning a system from a ‘high’ to ‘low’ stake approach to accountability.

3.3 The Context

Since devolution in 1999, Wales has had responsibility for its own education policy, which has gradually diverged from the other countries in the UK, particularly in its rejection of England’s more market driven reforms. The Welsh alternative has combined a complex mix of collaboration and partnership working with the gradual adoption of aspects of high stakes accountability measures.

A key moment in the development of the system occurred in 2009, when Wales was ranked the lowest of all the countries in the PISA rankings of the UK, being graded 43rd in mathematics; 41st in reading; and 36th in science out of the 65 countries who were included in the survey. This, in part, resulted in then Education Minister, Leighton Andrews, initiating a raft of reforms to the education system, including the amalgamation of the school improvement services of the 22 local authorities into a four regional consortia (Andrews 2011a, b). Andrews was bullish in his pronouncements around leveraging accountability into the system and looked to increasingly hold schools, and local systems, to account for schools' perceived under performance (Andrews 2011a).

The four regional education consortia worked with central government on a National Model for Regional Working, which represented a transformation of the middle tier of the Welsh education system. The proposed model was radical in that it envisaged a more collaborative partnership-based approach, for an education system that had been managed and evaluated largely through traditional hierarchical arrangements. The Auditor General for Wales was initially pessimistic about the extent to which the Welsh Government, consortia and local authorities would develop collaborative relationships in which 'strengths, weaknesses, developments and problems are shared, and the best solutions sought.' (Auditor General for Wales 2015 p.32). It was within this context that the Central South Consortium launched its 'Challenge''.

In developing the Central South Wales Challenge, use was made of lessons from research and experiences elsewhere. Particularly influential was the conceptual work carried out by David Hargreaves in relation to the development of a self-improving system: capitalizing on the benefits of clusters of schools; adopting a local solutions approach; stimulating co-construction between schools; and expanding the concept of system leadership (Hargreaves 2010, 2011, 2012a, b). Further design influences were provided by the City Challenge programme, which took place in three English city regions (see Ainscow 2015, for a detailed account of this initiative).

The Central South Challenge strategy was to develop a 'self-improving school system' based on six 'principles',

- Schools are communities where collaborative enquiry is used to improve practice;
- Groupings of schools engage in joint practice development
- Where necessary, more intensive partnerships support schools facing difficulties;
- Families and Community organisations support the work of schools;
- Coordination of the system is provided by school leaders;
- Local authorities work together to act as the 'conscience of the system'. (Central South Consortium Business Plan 2016, p.3)

The initial design influences on the Central South Wales (CSW) Challenge were provided by the City Challenge programme. Applying the strategic lessons learnt to

Wales required a recognition that there existed a broad political rejection to more competitive 'free' market approaches, combined with a desire to create a unique 'Welsh approach' to challenging the link between social deprivation and under achievement (Welsh Government 2014).The strategy that emerged, overseen by a group made up mainly of local headteachers, resulted in the launch of multiple collaborative school improvement initiatives, see Figure 3.1, that quickly engaged school leaders in a number of forms of collaborative learning and school improvement.

3.4 The Central South Wales Challenge

Figure 3.1 sets out the four key strands of the Central South Wales Challenge as advertised to schools in the region, these were: School Improvement Groups of 6–8 schools, Hub schools that provided a range of professional development and

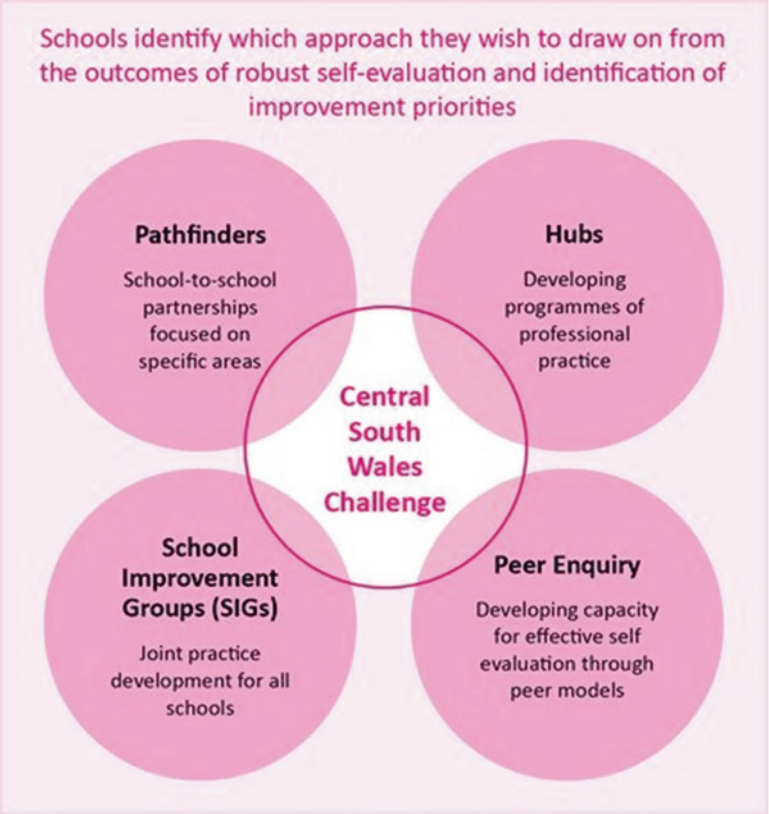

Schools identify which approach they wish to draw on from the outcomes of robust self-evaluation and identification of improvement priorities
Pathfinders
School-to-school partnerships focused on specific areas
Hubs
Developing programmes of professional practice
Central South Wales Challenge
School Improvement Groups (SIGs)
Joint practice development for all schools
Peer Enquiry
Developing capacity for effective self evaluation through peer models

Fig. 3.1 The main collaborative strands Central South Wales Challenge

professional learning opportunities, Pathfinder partnerships in which higher capacity schools supported more vulnerable ones, and the Peer Enquiry process.

All schools were placed in collaborative 'School Improvement Groups' (SIGs) of around 6–8 schools. The composition of the SIGs were designed to break down barriers between schools, and with this in mind each consisted of schools from across the five local authorities and with a mix of socio-economic intakes. Consideration was also given to schools' capacity to improve and so SIGs were composed of schools at different points on their improvement journeys. Meanwhile, some 40 odd schools with an understanding of enquiry were appointed as 'Hubs' and funded to engage in joint practice development with networks of local schools. There was also a number of 'Pathfinders', which were generally pairs of schools in which a higher capacity school provided focus support to the other school. The final strand of collaborative working was cohorts of headteachers being trained and supported to undertake 'Peer Enquiries' into each other's schools. The term 'enquiry' was used rather than 'review' to emphasize that the process was based upon mutual learning around a focus, rather than a process being 'done to' a school, it also distanced it from the language of external accountability systems.

The overall implementation strategy was to move from 'scope to depth', this meant initially focusing on the widespread mobilization of school leaders and involvement of practitioners in collaborative working and school-to-school support before gradually deepening the intensity and effectiveness of this way of working. The strategy was adopted in response to three key characteristics of the local system. Firstly, there existed a high degree of fragmentation within the system, arising from a lack of trust between local authorities and across its' different tiers (Evans 2015; Dixon 2016). The degree of fragmentation had suppressed previous attempts to develop collaborative ways of working so by rapidly expanding the number of schools and practitioners involved, the CSW Challenge hoped to increase levels of trust and establish supportive professional norms around collaborative working.

Secondly, historically low levels of engagement in collaborative school improvement approaches, in comparison with other parts of the UK due to policy and funding differences, meant there was a lack of school leaders with a substantive knowledge of school-to-school based enquiry methods. The Challenge therefore needed to offer a range of collaborative approaches that encouraged school leaders with very different levels of expertise to become engaged. Finally, competition between schools as a driver of improvement in Wales, was not as significant a policy lever as in other UK education systems, notably England, because of the Welsh government's opposition to quasi-market approaches. Existing patterns of competitive interactions were, therefore, less likely to be a major dampener on schools' willingness to participate in collaborative working.

The advantage of adopting a 'scope-to-depth' strategy was that it brought about both cultural and structural change. Culturally, it gradually reconfigured school leaders' understanding of their role within a self-improving system by normalising collaborative working. Structurally, both the professional learning landscape and approach to school improvement, previously dominated by centrally designed and delivered training programs and local authority employed advisors, was gradually replaced by school-to-school collaborative working.

3.5 System 'Layering': From Categorisation to Peer Review

As a whole, the Challenge adopted a range of strategies to transform existing school improvement and professional learning processes and structures. The introduction of Hub schools led to a rapid displacement of existing professional development programmes, previously provided by local authority staff, with school-based programmes. The Pathfinders were a form of 'bricolage' in which existing elements of intensive support for struggling schools were recombined into new more collaborative and reciprocal school-to-school approaches. Similarly, SIGs brought together school leaders into new professional networks, that took account of their schools' varying capacity for improvement, in order to provide support structures and encourage them to collaborate on school improvement initiatives with non-competitor schools.

The peer enquiry process was somewhat different from the other three strands in that it had to be developed whilst the National School Categorisation System, introduced in 2014 to replace a widely critiqued banding system, was still running. Annually schools in Wales are evaluated by Challenge Advisors (CA) from the consortia and placed into one of four support categories: green, yellow, amber or red, the latter being schools who require the most intensive support. Judgments are based on a number of pupil outcome measures and the schools perceived capacity to improve, using a national framework, in a process in which the CA also assesses the accuracy of the school self-evaluation.

'Categorisation', as it became known, increased the external pressure on schools to improve. Although there are no performance league tables in Wales, a school's categorisation is published annually on the My Local School Website, and discussed widely in the national media. In this way, in combination with rolling cycles of inspections by Estyn, the national inspection service, the process of categorisation forms a key part of the high stakes accountability system. School leaders are acutely aware of the impact of their professional status if their school has deemed to have 'dropped' a category. Similarly, consortia and local authority officials are held accountable for the profile of schools in each category and the overall trend for schools entering or leaving the high support categories.

The categorisation process was intended to support leaders of all schools access the support they required, in part brokered by the CAs. However, as the local system became more collaborative, some school leaders began to query its value as a means of acquiring the kind of support they required. School leaders from 'green' schools, in particular, increasingly felt that CAs either did not have the expertise or experience to add to their own self-evaluations, or were unable to broker-in the kind of support that would help them to improve their school. It was from within this group of school leaders that the idea of developing a different form of 'support' emerged, one organized around and accessed by a peer enquiry process.

In terms of the existing accountability system the peer enquiry process represented potentially the most subversive of the new collaborative strands. The original design of the process not only excluded the CAs responsible for a key aspect of the existing high stakes approach but also raised the possibility of replacing them in these roles, and of the categorisation system itself being radically reformed.

As an approach to system change the peer enquiry process was treated as an example of 'layering', in which new arrangements are placed on top of pre-existing structures (van der Heijden 2011). Layering involves the partial renegotiation of some elements of a system, while leaving others in place (Thelen 2003), because in part those involved in the change process lack the capacity to overturn the existing arrangements.

Layering is also an incremental approach to change, based upon gradually building up the 'thickness' of the alternative so that it becomes increasingly more valued within the system to the point where the pre-existing arrangements atrophy and are replaced by a process that has now become the new status quo. In this instance, the 'thickness' would be based on increasing the numbers of engaged school leaders, the development of more robust tools and procedures, during the phased development of the peer enquiry process, and most importantly of all, the development of layers of trust between school leaders.

3.6 The Phased Development of the Peer Enquiry Process

The peer enquiry process was driven forward by a small group of headteachers supported by officers from the consortium. The peer enquiry process progressed through four phases of development, beginning with a pilot involving some 6 reviews and 18 'Green' schools in 2014–15. This would gradually layer on top of the high stakes accountability process of categorisation, based on a national framework, a highly contextualised process based on reflexive professional dialogues with peers.

From its instigation, the term enquiry, not review or evaluation, was used and this reflected both the context in which was being developed and the aspirations of its' designers. The term peer enquiry was defined at the beginning of phase two as follows:

> 'The term Peer **Enquiry**, which connotes open suggestions or 'lines of enquiry' which are part of a continuous improvement system focused on practice is preferable to the stereotypical Peer **Review**, which may be static or one off. This is the term that is used in relation to the CSC model.
>
> The potential benefits for the Peer Enquiry Team members are also two way. Rigorous evaluation will offer chances to reflect on their own schools' provision in order to enhance or modify ideas, structures and initiatives 'back at base', as well as to offer professional development to the individuals themselves.' (CSC Peer Enquiry Phase 2, 2015 p.3)

The definition placed emphasis upon the on-going and reciprocal nature of the relationship between schools: an enquiry with another school rather than an evaluation being done to another school. Defining the process as mutually beneficial was seen as key in moving away from a 'sender-receiver' model of school-to-school support.

The more recent iterations of the model were based on a peer enquiry team of three who work with the 'host' headteacher who commissioned the enquiry. The team consisted of a lead enquirer, a headteacher with proven leadership expertise who had undertaken training in the peer enquiry processes and protocols. The second member would be a supporting headteacher, intending to undertake a peer enquiry in their own school. The third would be an associate member, a senior school leader aspiring to headship who would be either be affiliated to the lead enquirer's school or on a headship development programme.

The key responsibilities of the peer enquiry team were:

(a) To collaborate effectively with the host school
(b) To evaluate a specific focus or initiative requested by the host school
(c) To corroborate their pre-enquiry data and seek to confirm what the school says it knows and does
(d) To validate the school's best practices through a scrutiny of documentation, observation and discussion and to use collective professional judgment to identify and explore specific lines of enquiry that will enhance a school's self-evaluation processes
(e) To report back to the host school, offering clear and prioritised further lines of enquiry for the school to engage with, plus ideas for moving forward.

The relationship between the gradual accretion of the peer enquiry and atrophication of aspects of categorisation was illustrated by the changing role of the CAs. In the pilot scheme, they were kept at arms length from the process, as school leaders were concerned that their involvement would impact negatively on leaders' willingness to share openly their concerns. In phase two, the arms length approach continued but CAs role was now more defined and negotiated, but still peripheral. School leaders could invite CA to their personal feedback session and they would receive the same one page response to the final enquiry report that school governors received, which outlined 'strengths and areas of enquiry and the schools' response'. In addition, they were expected to broker support in line with the recommended lines of enquiry and, 'if required', link them to known best practice. In the same phase, CAs the time allocation for 'appraising Green schools' who took part in a peer enquiry was reduced.

In phase two, the wider roll-out of peer enquiry saw it being offered to all Green and Yellow schools, with some 33 reviews being undertaken. The consortium's role in the process, beyond funding release time for each enquiry team, focused on the training and support of new Lead Peer Enquirers and brokering schools into the enquiry process. In addition the consortium commissioned a report on the peer enquiry process (Matthews 2016) before the roll out of phase three to identify strengths and potential areas of development. This report, along with the on-going evaluation of the peer enquiry process carried out by the consortium, influenced its' development in two key areas.

Outline of the CSC Peer Enquiry Process
The peer enquiry process consists of four key stages:

Pre-Enquiry

The host headteacher sends a data pack to the lead enquirer 2 weeks before the enquiry visit with potential lines of enquiry. The lead enquirer spends half a day considering the data and opening up possible lines of enquiry.

The Enquiry visit

The enquiry team visits the host school over 2 days, during the first day:

- The lead enquirer meets team and briefs them re: data and lines of enquiry
- The enquiry team meet host SLT to discuss lines of enquiry
- Lesson observations, learning walks and interviews with stakeholders (including learners) (see toolkit for guidance on gathering evidence)

During day two the team engages in:

- Further enquiry – work scrutiny, deeper observational focus on key areas, more interviews etc.
- The team meet to reflect on evidence gathered – agreement about lines of enquiry
- Final meeting with SLT to offer feedback and discuss likely lines of enquiry
- *Post visit*
- Lead enquirer writes report and sends it for quality assurance (QA) via the strategic lead of the peer enquiry strand
- After QA clearance, the lead creates informal opportunity for host headteacher to see draft report and discuss action plan
- Report formally given to host school within 10 working days of peer enquiry
- The host headteacher reflects on the report and shares it with their governing body. The headteacher also completes reflective QA response and sends to CSC within 5 working days. The school SLT make modifications to the SER / SIP.
- *Follow up enquiry*

When a school has hosted an enquiry it may then commission a follow-up enquiry in the following year to look at progress and impact in relation to the key lines of enquiry.

Firstly, in terms of quality assurance, the increased numbers of enquiries had brought in a greater range of head teachers and greater efforts were needed to ensure the process remained consistent with its original principles. The key practical responses were: an affirmation of the non-negotiables of the process in the guidance materials, more comprehensive training for lead enquirers that included models of good practice, a toolkit[2] to facilitate the accurate and consistent gathering of evidence, greater emphasis on referencing the underpinning evidence within reports, and the appointment of Senior Peer Enquirers to review all the enquiry reports.

Secondly, more attention was placed on follow-on support and 'those further lines of enquiry that would make the greatest difference to pupils progress and outcomes' (Matthews 2016 p.3) that emerged from the initial enquiry visits. On-going evaluation had highlighted that the attention given to the initial enquiry process had not been matched by the consideration given to how schools subsequently followed up on the further lines of enquiry, the areas requiring development, that had been identified and how they implemented and evaluated any changes. The practical changes here were to prioritise these further lines earlier on in the process and providing better coaching and support to host schools in implementing and evaluating any changes. To give structure to this support, enquiry teams were now encouraged to revisit host schools 3–6 months after their initial visit and for schools to commit to cycles of enquiries on annual or biannual basis.

For the consortium, the challenges were now mainly focused on how to take the peer enquiry process to scale and establish it as a key aspect of the region's approach to school improvement. The challenges in taking peer enquiry to scale were both cultural and technical. The cultural challenge was that although peer enquiry had been initiated by local headteachers its promotion and funding by the consortium meant it could, by association, be seen as yet another aspect of a high stakes accountability mechanism. The consortium needed to ensure that the process was 'owned' by local headteachers. The technical issues were mainly concerned with how to resource the programme as it grew in size and regularity, against a backdrop of reductions in consortium funding, the programme had to reduce its overall costs.

In order to sustain the programme at scale, during phase three (2016–17) a more reciprocal resourcing model was proposed in which schools were to make a greater contribution. This involved only the time of the Lead Enquirer being paid for by the consortium. In addition, the 'supporting' headteacher, had to agree to engage in an enquiry in another school before they could commission an enquiry of their own. The third associate member no longer had to be from either the host or lead school but could be selected by the team from another school who would release them to support the enquiry as part of their professional development.

During this phase, in order to increase uptake and generate a greater sense of the process being owned by them, school leaders were encouraged to consider how to adapt the process to meet a wider range of school improvement needs. They were also encouraged to think how it might be focused on the quality of provision for different groups of pupils or be adapted to support then at different stages in their school improvement cycle, including dealing with a recent Estyn inspections. By phase three, the consortium was beginning to recognize that what had attracted the ‘early adopters’ to the programme had to be tweaked in order to draw in those who were less engaged with the collaborative reform agenda.

---

[2] https://www.cscjes.org.uk/search?query=toolkit&strict=false&popupUri=%2FResource%2F270e8943-e640-458a-8975-0f0aa416725e

The peer enquiry process was also opened up to ‘Amber’ schools and this resulted in the role of CAs becoming more differentiated. In respect of Green and Yellow schools, their involvement was at the discretion of the school leaders and largely dependent upon the outcomes of the peer enquiry process. A categorization of ‘amber’ meant schools received up to 15 days of support brokered by the CA. If an ‘amber’ school undertook a peer enquiry, the CA was given a more proactive role in framing the enquiry, by discussing the school’s need with the Lead Enquirer, and were in brokering in support to pursue action plans and further lines of enquiry. The CAs though were still not considered as ‘insiders’ and took no part in the enquiry process carried out by the school leaders.

By Phase 4 (2017 onwards), peer enquiry was established as a bi-annual model and consisted of an initial two-day enquiry process with a half-day follow up visit from their lead enquirer a year later. The consortium, besides providing administrative support and brokering schools into the process, also funded a part-time Lead Enquirer to carry out quality assurance, collate and share good practice, and to help in identifying schools that could provide on-going support in specific areas.

3.7 Determining Impact

The impact of the peer enquiry programme needs to be considered at a number of levels, from its effect upon individual schools and shifts in the dynamics between school leaders, through to its impact upon existing accountability structures.

Assessing the global impact upon schools and pupil attainment is highly problematic, not least because the sample of schools involved were self-selecting and therefore potentially unrepresentative of schools in the region as a whole. Not only was it initially taken up only by leaders of Green and Yellow schools, they were also those who were most engaged with the collaborative reform agenda.

Of the 66 schools involved in phase two, nearly 77% of primary schools who hosted a peer enquiry improved their categorisation over the next 12 months, compared to a regional average of 57%, whilst 83% of secondary schools who had hosted an enquiry improved their categorisation compared with a regional average of 65%. Similar improvement trends can be argued for at the pupil level in respect of Key Performance indicators. In primary, 81% of those schools involved in phase two improved their key stage 2 core subject indicators against a regional average improvement of 61%. Whilst in secondary schools, 83% of schools improved their key stage 4 level 2 plus against a regional average of 70%.

Evidence of impact in classrooms, on the professional development of school leaders, and wider effects on organizational learning is contained within the numerous reports and observational notes created during initial and follow-up visits. The majority of these reports though are unavailable for review due to the strict confidentiality protocols that they were collected under. As part of the consortium's own evaluation of the process, some 45 head teachers were surveyed as they attended a training session on the enquiry process prior to engaging fully with it, and individual and focus group interviews held with 8 experienced headteachers who had either lead on or hosted and enquiry.

The survey responses indicated that headteachers viewed the enquiry process as having a high degree of face validity, in that they expected it to have an impact upon them and their schools. They also indicated that, within the confines of the enquiry process, heads were more willing to 'open their doors' and engage in a critical dialogue about their school and its performance. The interview data explored experienced headteachers' motivations for initially engaging in the process and revealed how these were largely determined by their current level of engagement in collaborative working with other schools. For those already working in a collaborative structure, such as a SIG, it was primarily a means of deepening their existing work; whilst for those who were less engaged, it was an opportunity to widen their professional network by establishing an on-going relationship with another school leader.

The particular outcomes headteachers sought from their engagement fell into two broad headings that indicated the potential of peer enquiry to change the dynamics between them and how they related to accountability structures. Firstly, the headteachers sought support for their own strategies and approaches, particularly in respect of their identification of their school's challenges and how these should respond to them. The elements of the peer enquiry process that were key to these outcomes were: the value placed upon the 'empathetic' understanding of fellow headteachers; the ability to engage in open and frank dialogues around data to define the enquiry focus and to explore issues outside the prescriptions of categorisation or inspection; and the willingness of heads to accept critical feedback from colleagues and peers.

Secondly, headteachers sought a range of professional and organisational learning outcomes based on their own and their staff's engagement in the process. The key outcomes were associated with nature of the process, from the opportunities to engage in a data driven enquiry on an aspect of practice to evaluating the implementation of proposed changes. There was also a recognition, especially from those who were Lead Enquirers, that it presented an opportunity to challenge staff assumptions, often constrained by their limited experiences of very different types of schools and communities, of the range of challenges being faced by school leaders and how these might be addressed.

The impacts at the regional system level for the consortium were based on the scope and depth of engagement by school leaders, and any impacts it had had upon the local leadership 'culture'. By phase four, just over half of the nearly 400 plus schools in the consortium had been part of a peer enquiry.

In headteachers' accounts of their engagement in the process, they repeatedly emphasized the need for 'trust' in the process, particularly the individuals to whom they 'opened their doors'. Where school leaders' 'trust' in the peer enquiry process was reciprocated and replicated in different aspects of their engagement with other schools, for example in the nature of the on-going support they received, this developed further layers of trust. In addition, this increased the 'bandwidth' of their relationships with other school leaders, in the sense that they began to work collaboratively in other areas and aspects of school improvement.

Initially, leaders' 'trust' in the peer enquiry process was somewhat undifferentiated and impressionistic. That is to say, they trusted what they knew of the process, or at least what they had heard about it, and they might 'trust' those involved on the basis of their reputations as successful school leaders in the region, rather than because of any direct engagement with them. Leaders' engagement in the process also 'deepened' their trust in both the process and colleagues, in that it became both more defined and multi-faceted. They began to trust the specific assessments of colleagues, they began to trust in the value of pursuing new lines of enquires and the judgments made about their impact, and most importantly they began to 'trust' more in their own judgments and their and their staff's responses to the challenges they faced. The enquiry process had the potential to increase both knowledge and competency based forms of trust (Tschannen-Moran and Hoy 2000) between different school leaders, and leaders and their staff. Competence based trust is particularly powerful in overcoming the professional rivalries and tensions that often arise in contexts marked by competitions, as noted by Tschannen-Moran and Hoy (2000).

> In a study of CEOs and their subordinates, Gabarro (1978) found that friendship often followed the establishment of trust. In some cases, however, trust based on competence, good judgment, or reliability was maintained in spite of personal dislike. For example, it is quite possible for a school principal to dislike a teacher personally but trust his or her professional competence. (p. 560)

The initial evaluation of the peer review process emphasized that one of its key strengths was the extent to which it had 'the effect of reducing a strong dependency culture that has constrained the autonomy and outlook of school leaders.' (Matthews 2016 p.2). The cultural shift being observed was brought about by heightened levels of trust between local headteachers' which, along side greater mutual understanding of each others' schools and their capacities, increased their sense of collective efficacy. In the sense that the peer review process helped established a collective sense that not only should they support other schools but that they had the ability to make a positive difference to each others' schools.

For the consortium, it was this layering of different forms of trust, and its impact upon the culture amongst school leaders that, as much as wide spread engagement and the quality of the process, 'thickened' the peer enquiry process and made it an increasingly valued part of the local approach to school improvement. However, at the point of writing, the impact of layering the peer enquiry process on the national accountability structures, now described as the evaluation and improvement arrangement, is unclear. New arrangements are currently being developed and proposals and ideas floated, including the 'evolution' of the current school categorisation model. Such an evolution might involve an increased reliance on a renewed and more expansive school self-evaluation processes, the use of CAs to selectively review only a sample of these evaluations, and what has been described as peer engagement in the self-evaluation process. Over time the government has rolled back from describing this as peer review process, and developing a national model, faced with concerns over the workload implications for headteachers and the difficulties associated with imposing 'horizontal' accountable structures.

3.8 Power, Politics and System Reform

Why, then, is the study of a middle tier reform such as the development of the peer enquiry in the Central South consortium important? In part because it provides an alternative narrative to the rather crude construction of system reforms as either top-down or bottom up, and gives some insights as to how the political context and issues of power can play out within an incremental change processes. It is an account of how groups of practitioners when they experienced an increasing discrepancy between the 'high stakes' associated with an existing hierarchical accountability structure and the support it offered them were able to gradually develop a new 'horizontal' process. Although lacking the power to remove the existing centrally imposed system the consortium were able to encourage the 'layering' on top of an alternative more collaborative and reciprocal approach.

The phased development of the enquiry process also provides insights into the way in which layering can, if approached with a degree of subtlety, introduce change gradually in a way that avoids disrupting or directly challenging vested interests. The vested interest in this case were the Challenge Advisors, whose power was derived from their local control of a key element of a centrally enforced high stakes accountability structure: categorisation. The phased development although resisted by CAs did not provoke a direct counter-movement by these 'upholders of the status quo' (Streeck and Thelen 2005) and gradually supported a reconfiguration of their role.

The case adds to the literature on 'layering' by illustrating the importance of a key layer, that of the interactional dynamics, the degree of trust present and 'bandwidth', of the relationships between key players in a system. This layer is particularly important when considering how to move from high to low stakes accountability structures because of the recognized negative impact of the former on school leaders, and practitioners, professional identities, sense of autonomy and collective efficacy.

The peer enquiry process affected the power structures within the system through changing the dynamics within the professional networks of local headteachers. It initially drew together leaders around their dissatisfaction with an existing accountability existing system but through their commitment to a collaborative peer enquiry process it developed a shared 'identity'. An identity in which they saw themselves as part of the same 'self' who were collectively responsible for helping improve the local system.

References

Ainscow, M. (2015). *Towards self-improving school systems: Lessons from a city challenge*. London: Routledge.

Ainscow, M., Chapman, C., & Hadfield, M. (2019). *Changing education systems*. Abingdon: Routledge.

Andrews, L. (2011a). *Teaching makes a difference*. Cardiff: Welsh Government. (Speech at the Reardon Smith Lecture Theatre, Cardiff, 2 February) [Online]. Available at: http://wales.gov.uk/topics/educationandskills/allsectorpolicies/ourevents/teachingmakesadifference/;jsessionid=zDkvNBbKs4NZHXYzTJswvTRTVpnjTx1SGVt6BXGBkC2PYQQ7Zv1b!1895062788?lang=en. Accessed 1 July 2015.

Andrews, L. (2011b). *Raising school standards*. Cardiff: Welsh Government (Speech at Cardiff University, 29 June) [Online]. Available at: http://wales.gov.uk/topics/educationandskills/schoolshome/raisingschoolstandards/?lang=en. Accessed on 1 July 2015.

Auditor General for Wales. (2015). *Achieving improvement in support to schools through regional education consortia – an early view*. Cardiff: Auditor General for Wales.

Central South Consortium Joint Education Service. (2016). *Central South Consortium Business Plan.* Abercynon: Central South Consortium Joint Education Service.

DfES. (2014). *Qualified for life: An education improvement plan*. Cardiff: Welsh Government.

Dixon, P. (2016). *Testing Times: Success, Failure and Fiasco in Welsh Education Policy Since Devolution*. Cardiff: Welsh Academic Press.

Evans, G. (2015). *A class apart: Learning the lessons of education in post-devolution Wales.* Cardiff: Welsh Academic Press.

Hargreaves, D. H. (2010). *Creating a self-improving school system.* In *Nottingham*. National: College for School Leadership.

Hargreaves, D. H. (2011). *Leading a self-improving school system*. Nottingham: National College for School Leadership.

Hargreaves, D. H. (2012a). *A self-improving school system in international context*. Nottingham: National College for School Leadership.

Hargreaves, D. H. (2012b). *A self-improving school system: Towards maturity*. Nottingham: National College for School Leadership.

Mahoney, J., & Thelen, K. (2011). A theory of gradual institutional change. In J. Mahoney & K. Thelen (Eds.), *Explaining institutional change: ambiguity, agency and power* (pp. 1–37). New York: Cambridge University Press.

Matthews, P. (2016). *Peer enquiry in schools in the central south central consortium: a review of progress*. Ty Dysgu: Central South Consortium.

OECD. (2014). *Improving schools in Wales: An OECD perspective*. Paris: OECD Publishing.

Streeck, W., & Thelen, K. (2005). Institutional changes in advanced political economies. In W. Streeck & K. Thelen (Eds.), *Beyond continuity: Institutional change in advanced political economies* (pp. 1–39). Oxford, Oxford University Press.

Suggett, D. (2014). *Networking as system policy, balancing vertical and horizontal dimensions*. Paris: OECD CERI.

Thelen, K. (2003). How institutions evolve: Insights from Comparative-historical analysis. In J. Mahoney & D. Rueschemeyer (Eds.), *Comparative historical analysis in the social sciences* (pp. 208–240). New York: Cambridge University Press.

Tschannen-Moran, M., & Hoy, W. K. (2000). A multidisciplinary analysis of the nature, meaning, and measurement of trust. *Review of Educational Research, 70*(4), 547–593.

van der Heijden, J. (2011). Institutional Layering: A review of the use of the concept. *Politics, 31*(1), 9–18.

Mark Hadfield Over the last 20 years Mark has written and researched extensively on the leadership of school and system level improvement efforts. He has advised and supported national and local governments in developing collaborative approaches to school improvement and professional learning. Mark has been centrally involved in setting up a range of school networks both nationally and internationally.

Mel Ainscow CBE is Emeritus Professor of Education at the University of Manchester. A long-term consultant to UNESCO, he is currently working on efforts to promote equity and inclusion globally. Examples of his writing can be found in: *'Struggles for equity in education: The selected works of Mel Ainscow'* (Routledge World Library of Educationalists series, 2015).

Part III

Peer Reviews in the Context of High Stakes' Accountability. Intended and Unintended Consequences

Chapter 4
Self-Policing or Self-Improving?: Analysing Peer Reviews Between Schools in England Through the Lens of Isomorphism

Toby Greany

Abstract Peer reviews are not compulsory for schools in England, but they have become increasingly common in recent years. There is no single model for how peer reviews operate, but they generally involve staff from at least one other school in reviewing practice in the host school and feeding back their findings. This chapter reviews case study examples and data drawn from two recent studies led by the author (Greany T, Higham R. Hierarchy, markets and networks: analysing the 'self-improving school-led system' agenda in England and the implications for schools. IOE Press, London, 2018; Greany T. Sustainable improvement in multi-school groups. Department for Education, London, 2018). It analyses this evidence to assess whether and how peer review reflects the three forms of isomorphism (coercive, mimetic and normative) identified by DiMaggio and Powell (The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. American Sociological Review, 48:147–160, 1983). It finds evidence for all three forms of isomorphism, although levels of normative isomorphism vary and depend on the values and interests of different leaders. This analysis supports Greany and Higham's argument that peer review reflects a level of self-policing by schools in response to England's hierarchical and panoptic accountability system. The more recent study (Greany T. Sustainable improvement in multi-school groups. Department for Education, London, 2018) indicates that many Multi-Academy Trusts (MATs) are moving away from pure peer review models towards more hierarchically controlled approaches to assessing school quality. The chapter concludes by discussing these findings in relation to wider developments in the English school system as well as debates around quality, innovation and homogenisation in contemporary school systems.

T. Greany (✉)
University of Nottingham, Nottingham, UK
e-mail: toby.greany@nottingham.ac.uk

© Springer Nature Switzerland AG 2020

D. Godfrey (ed.), *School Peer Review for Educational Improvement and Accountability*, Accountability and Educational Improvement,
https://doi.org/10.1007/978-3-030-48130-8_4

4.1 Introduction

Few would argue with the assertion that publicly funded schools should be high quality (Woods and Macfarlane 2017), so the more significant questions are around what is meant by quality and how it can best be secured across multiple schools serving a range of different contexts. These questions raise further issues: does a focus on quality inevitably require some degree of standardisation and consistency, for example in the ways that school outcomes are assessed and the ways that schools are held accountable for these outcomes? If so, will such standardisation stifle innovation by imposing restrictive homogeneity, or will it enhance innovation by focussing collective improvement efforts on clearly defined and widely shared goals and success measures? Finally, what kinds of governance, accountability and support mechanisms might be most appropriate in order to balance the need for both quality and innovation in school systems?

School reform and governance changes across different school systems around the world have grappled with these issues over several decades (Greany 2016). However, in recent years a degree of consensus appears to have been reached: for example, both the OECD and the World Bank now argue that policy makers should grant schools a level of autonomy, especially in relation to curriculum and pedagogy-related matters, whilst holding them accountable for clearly defined outcome and quality measures (OECD 2013; Bruns et al. 2011).

School quality has thus become equated with accountability processes that can measure and benchmark schools using standardised data (Ozga 2009). Yet, in the process, it seems that innovation, agency and adaptive responses to different contextual needs can become compromised. For example, Hallgarten et al. argue that modern education systems are stuck 'in improvement mode' (2015), with a narrow focus on raising measurable standards preventing schools from focussing on the real needs of learners in the twenty-first century. These commentators argue that, as a result, many school systems face a crisis of legitimacy, with issues such as learner disengagement and stress, growing costs, frustrated teachers, challenges with equity and a mismatch with society's real needs (Sahlberg 2010).

Addressing these issues is not straightforward. Having simple and widely understood ways to evaluate school quality and to hold school leaders accountable for the use of public funds is arguably important, and can enhance the legitimacy of a system in the eyes of parents, employers and other stakeholders (Ehren et al. 2015a, b). Such approaches can also enhance equity, for example by ensuring that minimum standards are achieved and that additional resources can be focussed on children or schools where this is not the case (OECD 2018). But the downside may be that setting out and enforcing hierarchically defined standards for school quality will lead to homogeneity and inflexibility at a time when schools and education systems should arguably be more rather than less responsive to the changing needs of students and the changing world that they are growing up in (OECD 2015a, b).

Some school systems have attempted to resolve this by implementing ambitious reforms from the centre aimed at securing key innovations, for example to develop

'21st Century skills' in the curriculum. However, these centrally driven reforms tend to have limited impact at classroom level (Hall 2013) and some have been actively rejected by parents and wider stakeholders (Newton and Da Costa 2016; Waslander 2010). Other systems have introduced market reforms aimed at increasing choice and making schools more responsive to their communities, but the evidence of success is limited and such diversity can impact negatively on equity (Lubienski 2009).

These examples reinforce the message from O'Leary and Craig, who argue that 'central prescription takes us only so far and decentralised policies don't take us very far at all' (2007:8 cited in Cousin 2019:7). This recognition that both hierarchical and market-based forms of co-ordination have their limitations has led to a growing interest in network-based approaches to co-ordinating school systems in recent years (Suggett 2014; Hargreaves 2012).

Peer reviews between schools in England offer an interesting lens on these issues (see Box 4.1 for a summary of what peer reviews involve and how they have developed in England). Peer reviews undoubtedly represent an innovation, at least in the process of how schools work together to evaluate quality, but to what extent do they reflect a more fundamental shift in the English school system, away from hierarchical and market-based forms of control and towards network-based forms of governance (Ehren and Perryman 2017)?

Proponents of peer review argue they represent a move away from hierarchical forms of accountability, inspection and performance management of schools (Matthews and Ehren 2017; Berwick and John 2017; Gilbert 2012), opening up the potential for polycentric models of accountability that balance the perspectives of different stakeholders and encourage the development of collective effort and efficacy (Janssens and Ehren 2016). Others highlight the potential for peer reviews to enable schools and teachers to learn from and support each other, allowing effective practices to spread across a network (Matthews and Headon 2015). Hargreaves (2012) argues that such rigorous peer evaluation and challenge is the basis for a self-improving school system.

However, from a critical perspective, peer reviews may simply indicate the next stage in the evolution of the hierarchical accountability system, with schools self-policing their own performance in order to conform to the requirements of the performance management framework (Greany and Higham 2018). Such an analysis reflects the argument that school inspections form part of a wider panoptic – or post-panoptic – regime in the English context (Courtney 2016; Perryman 2009). Thus, schools face constant surveillance and the potential for punitive sanctions if they are deemed to be failing, so they must work to internalise and perform to the standards and expectations set by the inspectorate and accountability regime, even though these standards are frequently fuzzy and are constantly changing. In this interpretation, peer reviews, like school self-evaluations, reflect an internalisation of these requirements by schools and an attempt by schools to conform to the externally established standards they have been set.

This chapter debates these issues by drawing on case study examples and data from two recent studies of school networks and partnerships led by the author

(Greany and Higham 2018; Greany 2018).[1] The central question it seeks to address is whether peer reviews serve to reinforce the external accountability system and quality metrics, in the process making schools more homogenous, or whether peer review offers a means for schools to take ownership of what is meant by 'quality', enabling diverse, innovative responses to contemporary challenges? It analyses these examples through the lens of new institutional theory, in particular the three forms of isomorphism (coercive, mimetic and normative) articulated by DiMaggio and Powell (1983).

Box 4.1: Peer Reviews Between Schools – A Recent Phenomenon in England

Peer reviews between schools in England are a recent phenomenon, but have become increasingly common in recent years (Matthews and Headon 2015). For example, an extensive review of the school leadership landscape in England published in 2012 made no mention of peer reviews (Earley et al. 2012), whereas a national survey of primary and secondary headteachers published in 2018 indicated that nearly half (44%) of all schools had engaged in peer review in the previous year (Greany and Higham 2018). The survey indicated that peer review is now one of the most common forms of improvement support for schools and that it is rated highly by school leaders in terms of impact (see Fig. 4.1). This increase has occurred despite the fact that peer review is voluntary for schools; while England's inspectorate (Ofsted) has required schools to undertake self-evaluations since the early 2000s, and uses these as part of its own external inspections, it has never required or encouraged peer review.

A number of organisations have developed different models of peer review, which are offered more or less commercially for schools and school networks to use. Examples include Challenge Partners, CUREE, the Education

(continued)

---

[1] Greany and Higham (2018) included four detailed locality case studies as well as a national survey and different statistical analyses. Greany (2018) included 31 case studies of school groups as well as a national survey. See Greany and Higham 2018 and Greany 2018 for detailed research methodologies.

The author is also part of a team that is currently evaluating the impact of a collaborative peer review model (School Partnership Programme) for the Education Endowment Foundation, although no data from that project is drawn on here. See https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/schools-partnership-programme-spp/ accessed 18.3.19.

Box 4.1 (continued)

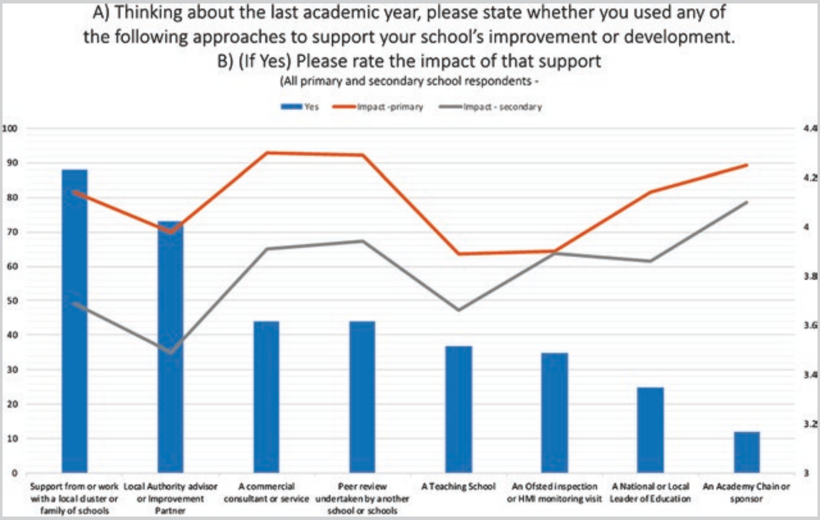


Fig. 4.1 Sources of improvement support for primary and secondary schools in England and their perceived impact by phase, from Greany and Higham 2018

Development Trust (EDT), the National Association of Head Teachers (NAHT), UCL Institute of Education and Whole Education.[2] However, the evidence presented in this chapter indicates that many – perhaps most – schools and school groups develop their own, bespoke approaches to peer review.[3] The analysis in this chapter relates to these bespoke approaches, rather than to any of the models listed above, except where a specific model is named.

This diversity of approaches to peer review makes it difficult to generalise about common features or principles underpinning the concept, although several illustrative examples are given throughout this chapter. What is clear is that peer reviews always involve staff from at least one other school in formally reviewing an aspect of practice within the host school and in

(continued)

---

[2] For details of the Challenge Partners, the Education Development Trust (EDT), the National Association of Head Teachers (NAHT), UCL Institute of Education models see Chaps. 5, 8, 9 and 11 of this book. For details on CUREE's model see http://www.curee.co.uk/peer_review and for Whole Education see http://www.wholeeducation.org/pages/overview/peoples_stories/763,0/peer_review.html accessed 4.1.18.

[3] For example, of the 47 case study schools visited by Greany and Higham, around half were involved in some form of peer review, but all of these were bespoke and none drew on the organisations listed here.

Box 4.1 (continued)
feeding back their findings. In this respect, peer reviews differ from other forms of school evaluation and accountability in England, such as internal/self-evaluations, formal inspections by Ofsted, and reviews undertaken by school oversight bodies such as Local Authorities or Multi-Academy Trusts. Equally, peer reviews differ from other forms of networking and collaboration between schools, for example through subject networks or school to school support, although such activities often occur alongside or as a result of peer reviews.

4.2 Hierarchical Accountability as a Driver of Behaviour in England's 'Self-Improving, School-Led System'

Understanding why and how peer reviews have developed in England in recent years requires an understanding of the wider shifts in policy and practice that have been underway since the election of a Conservative-led coalition government in 2010. That government and the Conservative governments that have followed it have pursued a set of policies that have included a rapid expansion in the proportion of schools that are academies, funded and overseen by central rather than local government and with increased 'freedoms' (i.e. autonomy) compared to Local Authority (LA) maintained schools.

A key strand in the rhetoric of the government's reforms since 2010 is that they will lead to greater innovation. The government has sought to reduce bureaucracy and claims that it trusts the profession to make appropriate decisions by increasing school autonomy (DfE 2010). The government has also promoted diversity, by introducing new types of school, such as free schools and University Technical Colleges, arguing that these will 'drive innovation… and offer pupils and parents a new approach to education' thereby 'galvanising others to improve, especially in areas where parents are significantly dissatisfied' (ibid: 58, 78).

Greany and Higham's (2018) research evaluated the nature and impact of these reforms through the lens of governance theory (hierarchy, markets and networks). It included four locality case studies as well as a national survey and a set of statistical analyses. The research found that while the policy agenda since 2010 has emphasised the development of a 'self-improving, school-led system', in which schools operate in 'deep' partnerships and networks (Hargreaves 2012) to share knowledge and capacity and thereby 'self-improve', the reality has been more complex and, frequently, problematic. While the Government has argued that its reforms are 'moving control to the frontline', the research shows that this is a partial and idealised account, with a strengthened accountability framework and continuing market forces serving to constrain the professionalism of teachers, to limit the autonomy of schools and to shape the ways in which knowledge and expertise are codified and exchanged.

Greany and Higham (2018) illustrate the pervasive influence of the national accountability framework, particularly as exercised through the inspectorate Ofsted,[4] on the thinking and practices of schools. They characterize this influence in terms of 'constrained professionalism' and 'coercive autonomy'. This influence results from the impact of inspection grades on schools, with punitive sanctions and interventions for schools judged to be performing poorly as well as more subtle impacts on the status of schools within local competitive arenas as they compete for pupils, staff and resources.

The influence of Ofsted is evident in the focus of school self-evaluations, improvement planning and the commissioning of advice and practice inspections ('mocksteds') from Ofsted-savvy consultants. It is also apparent in the language and thinking of school leaders as they describe their work, particularly at secondary level, suggesting that the accountability requirements have become internalised and are driving a relentless focus on consistency of practice within schools.

Importantly, the accountability framework is associated with increasing stratification between schools. Greany and Higham (2018) analysed national data from school Ofsted inspections over a 10-year period, which showed a relationship between inspection grades and the changing socio-economic composition of a school's student body. Schools that sustained or improved their judgement to Outstanding in the 2010–2015 period saw, on average, a reduction in the percentage of students eligible for free school meals (FSM), while schools retaining or being downgraded to a Requires Improvement and Inadequate judgement saw, on average, an increase in FSM eligibility.

In this context, most schools in England have formed or joined networks, partnerships and alliances so that they can access information, support and challenge from other schools. Greany and Higham (2018) analyse these partnerships and find that they are usually seen to deliver benefits for their members, but that they are predominantly focused on meeting the demands of the accountability framework, as the examples of peer review set out below demonstrate. Forming and leading these networks is frequently problematic for schools in the context of market-based pressures to compete with other local schools. As a result, many of the partnerships analysed by Greany and Higham (2018) are more or less exclusive in their membership. Meanwhile, the government has encouraged and, at times, coerced schools to form or join Multi-Academy Trusts (MATs),[5] arguing that this structure will secure efficiency and effectiveness, although the statistical analysis undertaken for the research challenges this assertion.

---

[4] All schools in England are inspected by Ofsted and graded as either Outstanding, Good, Requires Improvement or Inadequate.

[5] Multi-Academy Trusts are charitable companies that oversee more than one academy, with a single board and CEO responsible for all aspects of performance and operations across the group. By July 2018 there were 1082 MATs overseeing 5850 academies in England.

Greany and Higham (2018) argue that peer reviews demonstrate the extent to which schools in England have internalized the accountability requirements and are now 'self-policing' their work. This chapter builds on and deepens their analysis of peer review, including by introducing additional examples from the research that were not included in the main report for reasons of space. It also updates that research by discussing evidence from Greany's (2018) more recent study of school improvement models in MATs, Federations (where a number of maintained **schools** come together under one governing body), Teaching School Alliances (TSAs, these are groups of schools with a designated Teaching School as its hub school that work together to support each other, provide teacher and leadership training and other functions) and Local Authorities (LAs, the local administrative bodies that oversee educational provision).

4.3 Isomorphism: 'What Makes Organisations So Similar?'

This chapter analyses peer review through the lens of new institutional theory, in particular the three forms of isomorphism (coercive, mimetic and normative) articulated by DiMaggio and Powell (1983). Their article builds on Max Weber's analysis of the ways in which rationality, manifested through bureaucracy, serves like an 'iron cage' to control and standardise human activity as well as on Giddens' theories on the structuration of organisational fields. This chapter assesses peer review in relation to the three types of isomorphism and asks whether it contributes to greater innovation or standardisation in the context of hybrid governance mechanisms.

DiMaggio and Powell argue that "bureaucratization and other forms of organisational change occur as the result of processes that make organisations more similar without necessarily making them more efficient" (1983:147). Whilst innovation does occur in the early stages of a new organisational field, such as state schooling, once the field becomes established, "there is an inexorable push towards homogenisation" (ibid: 148) as organisations seek legitimacy. They argue that this is because such "similarity can make it easier for organisations to transact with other organisations, to attract career minded staff, to be acknowledged as legitimate and reputable, and to fit into administrative categories that define eligibility for public and private grants and contracts". Critically though, "none of this, however, ensures that conformist organisations do what they do more efficiently than do their more deviant peers" (ibid: 153).

DiMaggio and Powell posit that homogenisation is a result of institutional isomorphism, defined as "a constraining process that forces one unit in a population to resemble other units that face the same set of environmental conditions" (ibid: 149). They set out three mechanisms for institutional isomorphic change, although these are not always distinct and can co-exist: (i) coercive isomorphism stems from political influence and the problem of legitimacy; (ii) mimetic isomorphism results from standard responses to uncertainty; while (iii) normative isomorphism, is associated with professionalization.

Ehren (2019) analyses the influence of school inspection systems on thinking and practice in schools through the lens of isomorphism. As yet, however, this analysis has not been undertaken in relation to peer review.

4.4 Peer Review as Coercive Isomorphism

Coercive isomorphism occurs in contexts where formal and informal pressures are exerted on organisations, either by other organizations upon which they are dependent or as a result of wider cultural expectations in society. Sometimes these pressures take the form of governmental mandates, particularly in a field such as publicly funded education, but they can also arise from the application of standard operating procedures in back office operating systems, such as contract law and financial accounting. These pressures may be felt as force, as persuasion, or as invitations to join in 'collusion' (DiMaggio and Powell 1983: 150).

The description of England's accountability system – with its floor targets, inspection handbooks, school rankings and powers of intervention – and the ways in which it influences the thinking and practices of schools is a clear example of coercive isomorphism. Ehren's (2019) analysis of school accountability systems explores this influence in detail.

Despite being voluntary, peer reviews can still indicate coercive isomorphism if schools use them to prepare for an Ofsted inspection or to legitimise a set of actions that will bring the school in line with the expectations in the accountability framework. Such collusive behaviours are reflected in Greany and Higham's (2018) characterisation of peer review as 'self-policing', a term that comes from one of the head teachers they interviewed.

Vignette 4.1, below, illustrates the ways in which the headteacher of a standalone secondary academy uses peer review to set a demanding agenda for change across the school. It is notable that he chooses to work with a government-designated National Support School, providing a level of officially approved legitimacy, and that the partnership is non-local, and so does not risk sharing knowledge with local competitor schools.

The vignette makes clear that peer review, at least for this school, is not regarded as an opportunity to develop collaborative work and learning between staff working at different levels. Rather, it is a leadership-level-only activity, which positions the school's teachers as the cause of the 'problem' (i.e. the 'flat-lining' in exam results) which must be addressed through greater prescription and tighter oversight. The actions resulting from the review are aimed at creating a 'no excuses' culture, including through the introduction of non-negotiables, quality assurance mechanisms and 'drop-in' lesson observations, all suggesting a move towards bureaucratic rule-bound performance management.

The fact that the peer review is more blunt than the School Improvement Partner (SIP) report, and so has to be toned down before it can be shared with staff, is also revealing. Rather than asking questions to encourage self-reflection, as the SIP does, the peer reviewers name specific staff who they say should be sacked. This idea that peer reviews allow for honest, even brutal, feedback recurs through many of the interviews with senior school leaders in the study and is also a theme in Vignettes 4.2 and 4.3. Several interviewees acknowledge that this 'honesty' can be demotivating for staff, for example if they receive such feedback directly. This relates to the points made below, in relation to Vignette 4.3, about the types of 'strong' school leaders who are prepared to engage in peer review.

DiMaggio and Powell make the point that isomorphism does not necessarily make organisations more effective. Greany and Higham's (2018) research design was not intended to enable an evaluation of the impact of peer review over time. However, it is notable that the school described in Vignette 4.1 has dropped steadily in its exam performance in the 3 years since the case study visit.[6] This might not be surprising given the wider literature on school leadership and improvement, which indicates that successful schools tend to be characterised by high trust, aspirational and professional cultures in which staff and students are continually learning, rather than by managerial, rule-bound performance management-focussed processes (Daly and Chrispeels 2008; Hopkins et al. 2014).

Vignette 4.1: The Stand-Alone Secondary Academy – Peer Review as Coercive Isomorphism

The headteacher of this stand-alone converter academy acknowledges that he does not seek significant partnerships with other local secondary schools because he sees them as competitors. For this reason, the headteacher has worked with two other secondary schools from further afield to develop a peer review model. One of the other schools is a National Support School (NSS).[7] Reviews take place annually in each school, undertaken by heads and senior leaders from the other two schools, who present a written report on their findings.

The first review at the case study school investigated the school's hypothesis that there was an issue with its evaluation of teaching quality, given that this had found 93% of teaching to be Good or Outstanding based on lesson observations, yet the school's GCSE results were 'flat-lining'.

The review included scrutiny of pupil work and assessed the school's systems, processes and policies for supporting teaching and learning.

(continued)

---

[6] The school went from being above average for Performance 8 and average for Progress 8 in 2016, to being below average for Performance 8 and well-below average for Progress 8 in 2017 and 2018.

[7] These schools are designated by the government on the basis of performance criteria and are given additional funding to provide support to lower performing schools. The headteachers of such schools are designated as National Leaders of Education.

Vignette 4.1 (continued)

In parallel with this peer review, the school's practice in this area was also reviewed by its School Improvement Partner (SIP).[8]

The headteacher explained that the peer review report was harder hitting than the SIP's:

> *They [i.e. the peer reviewers] came out basically and said, "Your middle leaders aren't aspirational at all. You've got a real problem amongst your staff making excuses for kids coming from such poor backgrounds." It was very blunt and I did have to slightly temper it before I put it out to all the staff… [the SIP's report] was saying the same thing in a much more delicate, but still pointed, way... What I did was basically marry the two reports up and present to governor's the common themes. We did get good triangulation. It was good to have it verified by a separate team.*
>
> Head teacher, secondary academy, Ofsted Good

This triangulation helped the headteacher to reinforce with staff the need for a cultural shift across the school aimed at developing a 'no excuses' culture, informed by processes in place in the NLE/NSS school that had undertaken the review. The action plan for creating this cultural shift included developing a set of non-negotiables with quality assurance measures to ensure rigorous checking of teaching quality. In addition, lesson observations moved to unannounced drop-ins, in place of the previous model where teachers had 24 h notice. There was also a shift in focus from targeting resources at Year 11 to targeting these at pupils from Year 7, with the development of pupil progress 'flight paths', involving more regular monitoring of progress backed by interventions where required.

4.5 Peer Review as Mimetic Isomorphism

Mimetic isomorphism occurs in contexts where there is uncertainty, leading organisations to model themselves on similar organisations in their field that they perceive to be more legitimate or successful (DiMaggio and Powell 1983: 152). Preferred organizational models can be diffused by consulting firms or other organisations that promote 'best practices'.

Ehren (2019) shows how these processes operate in the context of school accountability in England. For example, Ofsted's thematic reviews and research reports on specific areas of practice are read avidly by schools, while its efforts to dispel the various 'myths' that abound (such as the idea that there is a preferred Ofsted teaching style or 'lesson') have had only partial success.

---

[8] An experienced advisor that the school buys in to review and advise on areas of practice as part of the annual improvement planning process

Greany and Higham (2018) show how the context for schools became more uncertain after 2010. This uncertainty was a result of several, overlapping factors: significant changes were made to the national curriculum and assessment model, while schools had to become more self-reliant as traditional forms of 'free' advice, support and challenge were reduced. For example, the head teacher in Vignette 4.3 explains that, with the collapse of the LA, "we had to sort our own houses out really".

In this environment, Greany and Higham (2018) argue that knowledge and expertise around aspects of school improvement became a more important 'commodity' for schools. High-status 'system leader' schools, for example those designated by the government to lead Teaching School Alliances (TSAs) and as NLE/NSSs, frequently stepped in to provide this 'commodity', for example by offering paid-for professional development courses and advice to other schools in areas such as how to achieve a good Ofsted inspection grade. Greany and Higham (2018) argue that this 'new economy of knowledge' has incentivised a focus on the types of expertise that can most easily be codified and commoditized (as 'best practices') rather than on the Joint-Practice Development and learning processes advocated by Hargreaves (2012) as essential for a self-improving system.

Vignette 4.2, below, indicates the ways in which peer reviews can facilitate mimetic isomorphism. The head teacher faces high levels of uncertainty, given his school has been judged Inadequate by Ofsted and so faces being taken over by a MAT. He has sought out a group of "widely respected" (executive) heads from across the region, who have "worked at national level on things like curriculum" and who are presumably, in DiMaggio and Powell's terms, seen as "legitimate and successful" (1983: 152). The head is keen for his staff to actually see these models of effective teaching in action, so has taken them, en masse, to visit an NLE's school for a day. This approach differentiates this vignette from the last one, in that the headteacher is clearly keen to involve his staff in a process of learning from other schools. However, the peer review process itself involves only members of the leadership group from each of the three schools. Despite leaving the host head teacher like they have been "beaten up", the peer review is followed by "a massive wave of support" in terms of practical ideas and resources that the school can use as it works to mimic the practices observed in and advised on by the respected, high performing schools.

Interestingly, having performed well below average at the time of its Ofsted Inadequate judgement, 3 years after the case study visit the school scored well above average (compared to national levels) for its performance in reading, writing and maths at Key Stage 2. This suggests that where peer review is coupled with a serious effort to facilitate the learning of all staff as a means of securing mimetic isomorphism, it can lead to improvements. These issues are returned to below in the discussion of isomorphism, innovation and improvement.

This vignette focusses on a school that has been judged Inadequate by Ofsted, and so arguably has a strong imperative to mimic the practices it observes in higher performing schools. But mimetic isomorphism is also apparent in the other cases, where the hierarchical pressure to enhance performance is less acute. For example, the school in Vignette 4.3, below, is judged Good by Ofsted, but its headteacher explains in the quote below how he benchmarks his school against the performance of the Ofsted Outstanding schools in the partnership and aspires to reach their levels, providing a clear example of mimetic behaviour:

> *I think that's the first time I've gotten very clear understanding of what outstanding data looks like. And it's a different league to ours… And our challenge… is that we have to match that standard. We can go on about different cohorts or whatever else. It doesn't matter. We have to be as good as they are. So, our conversation there is how we set up our aspirations. How do we put in place support for our staff, so that we are all at that level, and that they're sustaining that level?*
>
> Headteacher, primary maintained, Ofsted Good

Vignette 4.2: The Inadequate Primary School – Peer Review as Mimetic Isomorphism

This primary school had been judged Inadequate by Ofsted and was in the process of forced academisation (i.e. being sponsored by a MAT) at the time of the case study visit. The Head teacher had arrived at the school, his second headship, soon after the Inadequate judgement. He and his Deputy head argued that the school's previous Head had let the school become too insular, with virtually no recruitment from beyond the existing staff and very little collaboration with other schools. The new Headteacher's approach was therefore to open the staff's eyes to new and different models of practice, both within the local cluster and more widely. For example, he had taken the entire staff to visit a nearby National Leader in Education's school:

> *I wanted to show them the vision of what I wanted to achieve and where I wanted to be… Every teacher went down there and observed good and outstanding practice, looked at how the learning was structured and the language of learning. .*
>
> Head teacher, Maintained Primary, Ofsted Inadequate

The head had also invited in a group of other heads from across the region to undertake a peer review of the school:

> *I got four heads peer reviewing the school… I invited them in because they were people who were widely respected… I needed that – maybe this isn't the right choice of words – but maybe kind of that level of brutality and honesty…. They were all people with 2 or 3 schools, people with 500+ kids, people who've worked at national level on things like curriculum or whatever else.*
>
> Head teacher, Maintained Primary, Ofsted Inadequate

The Deputy Head argued that this peer review '*was really really good, and constructive, because when they said something was poor, they then had a suggestion for how to improve it*'.

The headteacher is now engaged in a regular cycle of peer reviews with this group of heads. The format is that the heads arrive in the morning, before the school opens, and meet with the host headteacher. There is no set focus for the review and no protocols. The group tours wherever they like in the school

(continued)

Vignette 4.2 (continued)

for two and half hours, talking to pupils and staff who have been warned that the review is taking place. They then meet back together without the host head and review what they have seen before feeding back to the host:

> *And then you call the head back in and you feedback. But then the next bit's critical because the person by the end of it, feels like they've been absolutely beaten up. But probably by five o'clock that evening they'll have had ten, fifteen emails: 'Here's a teaching and learning plan you might want to look at', 'Here's something I've used', 'Come and look at this other school', 'Would you like some help on this?', 'I can broker this for you' – so it's a massive wave of support straight after.*
>
> Head teacher, Maintained Primary, Ofsted Inadequate

4.6 Peer Review as Normative Isomorphism

The third source of isomorphic change stems from professionalization, as those working in a particular field establish professional standards, entry requirements and networks that set normative standards for how things should be done. Adhering to these professional standards and expectations is seen to help build organisational legitimacy, for example where organisations require certain qualifications as an entry standard for appointments.

Ehren (2019) highlights the decision taken by Ofsted after 2010 to increase substantially the proportion of serving school leaders trained as inspectors.[9] These serving school leaders spend a number of days each year working with Ofsted's core team of HMI (Her Majesty's Inspectors) to undertake inspections. Ehren highlights this as one of the ways in which the accountability model facilitates normative isomorphism, not least because these Ofsted-trained school leaders then draw on their training to inform their work in their own schools.

Vignette 4.3 describes the tensions that occur between two groups of primary head teachers in one town: six of whom have volunteered to be trained as Ofsted inspectors and six that have not. The six Ofsted-trained head teachers propose adopting peer review across the cluster, but the non-Ofsted trained heads reject this. The Ofsted heads decide to go ahead anyway and, from this, develop a wider partnership and the SUCCESS TSA, which serves to entrench further the divide between the two groups.

This vignette reveals that normative isomorphic processes are not uniform and can involve significant tensions between different professional perspectives and value-sets. The head teacher explains that the six heads involved in SUCCESS "all viewed each other as equals," describing them as "fairly arrogant, strident characters." These six heads were all officially designated as system leaders, one as an NLE and the other five as Local Leaders of Education. The head teacher acknowledges that there has been an historic divide between the two groups, explaining that the other head teachers in the cluster saw the SUCCESS group as "class traitors" because they had been Ofsted trained.

---

[9] For details see: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/465626/Ofsteds_inspection_workforce_from_September_2015.pdf accessed 22.3.19.

The divide between the two groups of heads arose out of the choices and values of individual leaders, but these differences have become physically embodied in the local partnership structure. The Ofsted-trained heads felt there was a danger that their schools could slip back in performance terms after the LA collapsed, so they chose to use peer review as a means of "avoiding complacency". Meanwhile, the head teachers who resisted such inspectorial approaches are characterised as "vulnerable" "losers" in the new "capitalist" system.

This sense that there are two 'classes' of leader – those who embrace Ofsted and peer review and those who don't – comes through from the other examples in the study. For example, the head of the school in Vignette 4.2 explained: "You've kind of either got the stomach for it [i.e. peer review] or you haven't" and expressed sadness that "it's the people who often need it the most, who don't engage in it, because they feel scared or they're nervous or they're worried". This suggests that normative isomorphism is not a uniform process, at least in a system that is going a through a process of rapid change, but one that is dependent on individual character and professional values as well as circumstance. Some individuals, it seems, are simply more likely to volunteer to be trained as an Ofsted inspector, to apply to be designated as a system leader, and/or to see peer review as an opportunity for improvement of their school. Others reject such approaches, perhaps out of fear, but perhaps because they adhere to a different notion of what it means to be a leader.

Over time, it seems likely that the 'strong' style of leadership, as demonstrated by the SUCCESS heads, will come increasingly to dominate the system through a process that combines normative isomorphism with hierarchical incentives (Gronn 2002; O'Brien 2015). Essentially, the government encourages these kinds of behaviours and the leaders who demonstrate them are given greater influence over how the next generation of leaders should be identified and developed, meaning that, over time, their 'strong' style of leadership will come to be seen as the norm. For example, designated system leaders are funded and encouraged by the government to support or sponsor 'failing' schools, giving them the power to enforce their preferred ways of working and to appoint the next generation of head teachers. These same system leaders also tend to be invited to advise on new leadership standards for the profession and to design and deliver the government's National Professional Qualifications for Leadership (Cousin 2019).

Vignette 4.3: The SUCCESS Alliance – Peer Review as Normative Isomorphism

The headteacher of this primary school, which is located in a small town, considers that collaboration has become increasingly essential for school improvement, in particular because the formerly strong Local Authority declined rapidly after 2010, which meant "we had to sort our own houses out, really". However, he felt that the local cluster, which included 12 primary schools, had failed to recognise the implications of this. The key sticking point was when six of the primary school heads proposed developing a model of peer review. The other six primaries resisted this proposal, but the proposing group decided to do it anyway:

> *Literally, as soon as we mentioned doing inspections in each other's schools, the room just divided in two, from "over my dead body" to those which were, "fine"... which was why SUCCESS [TSA] formed, because we wanted to move things at a higher pace than some of the other heads.*
>
> Head teacher, primary maintained, Ofsted Good

The head teacher feels that peer review avoids complacency, for example after a successful Ofsted, as it identifies what still needs to be improved. Each school is entitled to a yearly review of one of its school improvement priorities by two SUCCESS colleagues. Reviews typically comprise: data analysis, lesson observations, work scrutinies and pupil conferencing. To date, the school has received two reviews, covering maths and punctuation and grammar.

The head teacher feels that the process has worked well and that this is due in part to established trust between, and confidence in, each other as partners:

> *I think, partly because we all viewed each other as equals. If I'm honest, we're fairly arrogant, strident characters who believe we're right… the headteachers that visited (my school) pulled no punches, telling staff what needed to improve, so that set the tone, if you like... (but) if we're saying we want our schools to improve, then we have to ask ourselves difficult questions.*
>
> Head teacher, primary maintained, Ofsted Good

Building on the peer reviews, the six schools had developed a range of wider partnership activities. One of the schools in the group had been designated as a Teaching School, with the other five schools as strategic partners in the SUCCESS Alliance. There is also discussion around the potential to become a MAT.

However, the head teacher acknowledges that the development of SUCCESS as a separate entity from the wider cluster has led to a division between the strong and less-strong schools in the locality:

> *SUCCESS appeared, because we felt we couldn't wait. The world was changing around us, and if we didn't do something, we'd be left on our own. I think it's unfortunate that probably the six strongest schools in [the cluster] formed SUCCESS. And that was to our shame, a little bit, I think, that the egalitarianism stopped. And I think that our vulnerable schools within [the cluster], within the locality, are on their own, because they weren't able or willing to join.*
>
> Head teacher, primary maintained, Ofsted Good

(continued)

Vignette 4.3 (continued)

The head of the school admits to feeling deeply conflicted by this development, but feels that his response was the only option in the context of the government's policy framework:

> *I think it's a capitalist model. It's about school-to-school competition, and the government's very hot on that, and for that, there are winners and losers. And right now, I've taken the pragmatic, yet morally dubious position of 'I want to be with the winners', and that means I have to leave out some losers, some people who are vulnerable, on the outside. And we know that they're there. We know that they'd bite our arm off to come and join us. But we can't have lots of voices in the room if we're going to move things on quickly. And that's not fair.*
>
> Head teacher, primary maintained, Ofsted Good

4.7 Recent Developments: A Move Away from Peer Reviews in MATs?

Greany's (2018) research provides an update on the Greany and Higham (2018) findings, both because the data was collected more recently and because it focuses on approaches within MATs, which have become a more central feature of the landscape in England in recent years.[10]

Greany and Higham's (2018) study included a statistical analysis of MAT impact as well as case study research in a small number of MATs. One of their findings was that MAT leaders feel a pressure to standardise practices across their member schools in the pursuit of higher standards, based on a perception that higher performing MATs are highly standardised. This suggests that coercive and mimetic isomorphism operate at MAT as well as school level, driven by the demands of the accountability framework and by the formal and informal messages promulgated by the government. Given this finding it is interesting to ask how peer reviews operate within MATs.

Greany (2018) identified a set of important contextual differences (such as size and composition) which influence how different MATs approach school improvement, indicating that this emerging field remains diverse at this early stage.[11] In terms of practices within MATs, the research found that most are focussed on standardising or aligning practices across member schools to some extent, although the extent of this differs between different areas of practice; for example, assessment practices tend to be more standardised, while curriculum and pedagogy tend to be less so.

---

[10] Greany and Higham's case study research was undertaken in 2015–2016, while Greany's was conducted in 2018.

[11] Interestingly, MATs themselves are now being encouraged and facilitated to engage in MAT to MAT peer reviews by organisations such as Challenge Partners and Education Development Trust. For example see: https://www.challengepartners.org/news/blog-evincing-mat-factor

The case study research found relatively few examples of peer reviews operating within MATs in the ways described in this chapter so far. Instead, it found that the majority of MATs were undertaking periodic school reviews led by a member of the MAT core team, such as the CEO, the School Improvement Director or, sometimes, an externally commissioned consultant. These MAT reviews generally take place termly or annually: although the regularity of these visits might be determined by an assessment of risk, with lower performing schools visited more often. The format of the MAT reviews is similar to a peer review, but with aspects that feel closer to a mock Ofsted inspection: for example, one School Improvement Director described how she identifies 'lines of enquiry' before she visits a school based on an analysis of the school's data and self-evaluation.[12] Overall, these reviews were clearly positioned as a means of securing hierarchical accountability (i.e. to the MAT) and of identifying any performance issues in schools. For example, one Executive Principal described these reviews as "peer reviews with extra rigour, as sometimes peer reviews tend to be a bit woolly or a love-in" (2018:170).

Several MATs described these reviews as 'peer reviews' and many of these did include staff from other schools in the process (i.e. alongside the core team members). Such involvement was seen to provide a formative and developmental process for the staff involved. It also allowed the staff to propose ideas and to provide support to the school being reviewed if needed. This willingness to get involved in helping collectively to address the issues identified suggests that, at least for some MATs, the focus is on both accountability and support.

4.8 Discussion

Di Maggio and Powell asked 'what makes organisations so similar?' and argued that the three forms of isomorphism operate separately and in tandem to drive this homogeneity. In many ways their implied critique of the bureaucratic 'iron cage' is one that the Conservative-led governments in power since 2010 would subscribe to. For example, shortly before he was elected as Prime Minister in 2010, David Cameron argued that "the era of big government has run its course". In its place he called for a 'Big Society' approach fit for a "post-bureaucratic age", in which "the model of state-run schools, accountable to ministers and education bureaucrats will be replaced by self-governing state schools accountable to parents" (Cameron 2009).[13]

In practice, as Greany and Higham (2018) show in detail, the changes introduced since 2010 have not unleashed the innovations that Cameron promised. The stripping away of Local Authorities combined with multiple changes to the curriculum, assessment and accountability systems has left schools feeling more sharply accountable in a system that has become increasingly centralised. The result is that schools feel more rather than less constrained and coerced, leading them to 'self-police' their work through voluntary peer reviews.

---

[12] Similarly, Ofsted inspectors develop 'lines of enquiry' based on an analysis of data.

[13] David Cameron, 'The Big Society', speech given on tenth November 2009 https://conservative-speeches.sayit.mysociety.org/speech/601246 accessed 10.5.19.

This suggests that Cameron's vision of a 'post-bureaucratic age' and of self-governing schools accountable only to parents was, at best, naïve. Rather, the state has defined, in particular via the Ofsted inspection framework, the features of what it deems a 'quality' school, and continues to enforce these features through its inspection and accountability system. This quality framework is used to judge and rank schools in a formal sense, but also works through the isomorphic processes described here to drive a level of consistency – or homogeneity – in how schools operate. DiMaggio and Powell argue that such processes are more particularly common in publicly funded and operated systems, such as schooling: "the greater the extent to which the organisations in a field transact with agencies of the state, the greater the extent of isomorphism in the field as a whole" (1983: 154).

The three vignettes outlined in this chapter illustrate the ways in which peer review facilitates these isomorphic processes, thereby revealing the ways in which schools voluntarily 'self-police' their work. The evidence of coercive and mimetic isomorphism is particularly clear: the school leaders in these vignettes use peer reviews to benchmark their own school against others that they see as more (or equally) legitimate at a time of significant uncertainty around how best to respond to changing policy and accountability requirements. They use the findings from these reviews to prepare for their next Ofsted inspection, largely by setting an agenda for change that focuses on emulating the structures and processes in place at the higher performing schools that undertook the review.

The evidence of normative isomorphism is also strong, although this process is still evolving in England's turbulent school landscape and there is some evidence of an alternative professional ethic that rejects peer review in Vignette 4.3. However, this professional dissonance should not be overplayed: in two of the four localities researched by Greany and Higham (2018), nearly all of the secondary schools were engaged in a collectively agreed approach to peer review, indicating a higher level of normative alignment – or, at least, a greater acceptance of the need to conform with the requirements of the accountability framework – than in the vignette. Greany and Higham (2018) note differences between the primary and secondary phases in this respect, with some primary leaders more committed to the model of LA scrutiny of schools (and therefore generally less likely to engage in peer review) compared with their secondary peers.

While the three forms of isomorphism are addressed separately here, in practice they overlap and interact, partly as a result of wider processes that result from hybrid governance processes. The vignettes frequently illustrate these interactions, but they also suggest a hierarchy in terms of how the three processes operate, with coercive isomorphism – expressed most simply as a fear of, and need to align with, Ofsted – acting as the primary driver of behaviour in all three examples. Greany and Higham's (2018) report shows in detail how the hierarchical pressure exerted by the accountability framework interacts with quasi-market pressures on schools, which result from the need to compete for students and resources within a context of parental choice and local status hierarchies. So, for example, school leaders report that by improving their school's Ofsted grade they can signal to parents that the school is successful, which can enable them to attract a more aspirational intake and thereby further enhance their likelihood of success in Ofsted terms. However, in the process, other local schools can be disadvantaged by these changes. These pressures serve to influence how and where schools collaborate: for example, we saw in Vignette 4.1 that the headteacher chose to collaborate beyond the locality due to competitive tensions, and in Vignette 4.3 the role of peer review in splitting the local community of schools into two groups.

Mimetic isomorphism can be seen as a parallel but, perhaps, less dominant process when compared with coercive isomorphism. The schools are motivated by – or are in collusion with – the coercive requirements of the accountability framework, but they must also engage in mimetic processes in order to identify and transfer systems, processes and practices from Ofsted-successful schools to their own in the context of considerable policy-generated uncertainty. However, the extent to which these mimetic processes lead to genuine changes in classroom practice is not straightforward. Because the peer reviews generally involve only one or two members from the senior leadership team in each school, there are limited opportunities for staff to learn from each other as part of the review process. As a result, the new processes introduced on the back of peer reviews may not have the desired impact: if anything, they may distract and demotivate teachers and undermine the kinds of high trust, collaborative professional cultures that are known to underpin school improvement. We saw this in the first vignette, where the school declined in its overall performance after the new practices were introduced. By contrast, the school in the second vignette, where all staff were involved in visiting and learning from the high performing school, has improved its performance. This suggests that peer review as mimetic isomorphism may have limited efficacy unless it is combined with a serious focus on enabling staff to learn from and to adopt and adapt the desired approaches to their context.

Meanwhile, normative processes overlay these developments, helping to explain which leaders are more and less likely to engage and to indicate why certain types of leadership behaviours become embedded across the system over time. What is less apparent in these brief examples, but comes through in Greany and Higham's (2018) full report, is how these forces interact to shape patterns of collaboration, competition and change across local school landscapes. This analysis reveals, for example, how national 'system leadership' designations, such as NLE/NSS status, can structure local governance arrangements and status hierarchies, determining which individual leaders hold most sway.

Turning to the question posed in the Introduction, it seems unarguable that the three examples of peer review described here are serving to reinforce England's high stakes, panoptic accountability system, through 'self-policing'. This challenges the arguments made by advocates of peer review, as outlined in the Introduction. Rather than the "self-accounting – even self-regulating school system" pictured by Matthews and Ehren (2017:50), the vignettes seem closer to Perryman's (2009:628) description of school self-evaluation: "The teachers (or rather management) are (now the) inspectors, but without the power to make judgements. They are merely the warder, not the director of the prison with the power to liberate or punish".

Greany and Higham (2018) do report that a minority of their interviewees saw peer review as providing "a space 'outside' the accountability system in which to be honest without fear of reprimand" (ibid:32). Certainly, there were some limited examples in the research that were less 'brutal' than the three vignettes, reflecting an ethos that could be seen as closer to 'appreciative enquiry'.[14] However, even these examples tended to be framed in relation to the accountability framework. For example, one head teacher described using peer review to identify and address the 'weaknesses' that she knew existed in her school within the relative privacy of a school to school partnership; however, she acknowledges that this work is still geared towards preparing for more formal monitoring visits and inspections.

Meanwhile, the evidence from Greany's (2018) more recent study of MATs indicates that these corporate school groups are moving away from peer review and towards models that can more rigorously hold school leaders to account for their performance. Greany's research shows that MATs are working in wider ways to develop standardised or aligned approaches to pedagogy, curriculum and assessment, based on a view that this will ensure efficiency and effectiveness. The corporate reviews described above are therefore serving to reinforce these wider processes of alignment and standardisation within MATs. Over time, as MATs come to further dominate the school landscape in England, it seems quite possible that these MAT corporate reviews will largely replace the voluntary peer review models described in this chapter.[15]

The final questions that this chapter has raised are less straightforward to answer empirically. Does peer review – and the wider school quality and accountability frameworks that drive it – really lead to the homogenisation of schools? Does having a consistent definition of school quality and a clear accountability framework ultimately support or hinder innovation? Certainly, some argue that adopting a shared definition of 'what makes a good school' across a system is helpful if it allows practitioners to understand where and how their own school needs to improve and if it provides a collective focus and shared language for improvement efforts (Ehren et al. 2015a, b; Bruns et al. 2011). From this perspective, homogenisation is to be welcomed – or at least is a price worth paying – if it means that all schools are adopting evidence-based practices and all children are benefitting from a minimum standard of education. In contrast, DiMaggio and Powell argue that homogenisation limits diversity and reduces levels of innovation. This argument chimes with the view, outlined in the Introduction, that many school systems are stuck 'in improvement mode' (Hallgarten et al. 2015) and are failing to adapt to the changing needs of children and societies. Proponents of this view argue that school systems need to reject "excessively bureaucratic models", arguing instead that "more organic metaphors and models might seem messy and unpredictable, but eco-systems and complexity have become the nature of the contemporary world" (OECD 2015a, b: 17; Greany 2019).

---

[14] See https://en.wikipedia.org/wiki/Appreciative_inquiry accessed 22.3.19.

[15] It is notable that several of the national organisations that have promoted peer reviews between schools are now developing MAT-to-MAT peer review models. For example, see: https://www.challengepartners.org/mat-peer-review accessed 13.5.19.

In practice, it seems that some systems are already developing a 'middle way' between these two perspectives. For example, inspection in the Netherlands is focussed at the network level (i.e. the school boards that are equivalent to MAT boards), as well as at school level, potentially allowing for a more polycentric and less hierarchical approach (Honingh et al. 2018; Janssens and Ehren 2016).

In conclusion, this chapter has provided original empirical evidence of peer reviews between schools in England. By analysing this evidence through the lens of isomorphism it has sought to provide an original and rigorous analysis, showing how peer review reflects all three isomorphic processes but particularly coercive forms in the context of England's panoptic accountability regime and quasi-market system. In discussing these findings the chapter raises important questions around the relationships between isomorphism, homogeneity, innovation and improvement in education in the context of hybrid governance. These themes arguably deserve further investigation through research across different international contexts and settings.

References

Berwick, G., & John, S. (2017). School networks, peer accountability and the brokerage of knowledge and expertise. In P. Earley & T. Greany (Eds.), *School leadership and education system reform*. London: Bloomsbury.

Bruns, B., Filmer, D., & Patrinos, H. A. (2011). *Making schools work: New evidence on accountability reforms*. Washington, DC: The World Bank.

Cameron, D. (2009). *The big society*. Speech available from https://conservative-speeches.sayit.mysociety.org/speech/601246. Accessed 10 May 19.

Courtney, S. (2016). Post-panopticism and school inspection in England. *British Journal of Sociology of Education, 37*(4), 623–642. https://doi.org/10.1080/01425692.2014.965806.

Cousin, S. (2019). *System leadership: Policy and practice in the English schools system*. London: Bloomsbury.

Daly, A., & Chrispeels, J. (2008). A question of trust: Predictive conditions for adaptive and technical leadership in educational contexts. *Leadership and Policy in Schools, 7*(1), 30–63.

Department for Education (DfE). (2010). *The importance of teaching: The schools white paper* (Cm 7980). London: Department for Education.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review, 48*, 147–160.

Earley, P., Higham, R., Allen, R., Allen, T., Howson, J., Nelson, R., Rawar, S., Lynch, S., Morton, L., Mehta, P., & Sims, D. (2012). *Review of the school leadership landscape*. Nottingham: National College for School Leadership.

Ehren, M. C. M. (2019). Accountability structures that support school self-evaluation, enquiry and learning. In D. Godfrey & C. Brown (Eds.), *An ecosystem for research-engaged schools: Reforming education through research*. London: Routledge.

Ehren, M. C. M., & Perryman, J. (2017). Accountability of school networks: Who is accountable to whom and for what? *Education Management Administration and Leadership., 46*, 942–959. https://doi.org/10.1177/1741143217717272.

Ehren, M., Perryman, J., & Shackleton, N. (2015a). Setting expectations for good education: How Dutch school inspections drive improvement. *School Effectiveness and School Improvement, 26*(2), 296–327. https://doi.org/10.1080/09243453.2014.936472.

Ehren, M. C. M., Gustafsson, J. E., Altrichter, H., Skedsmo, G., Kemethofer, D., & Hüber, S. (2015b). Comparing effects and side effects of different school inspection systems across Europe. *Comparative Education, 51*(3), 375–400.

Gilbert, C. (2012). *Towards a self-improving system: The role of school accountability*. Nottingham: NCSL.

Greany, T. (2016). Innovation is possible, it's just not easy: Improvement, innovation and legitimacy in England's autonomous and accountable school system. *Education Management Administration and Leadership, 46*, 65–85. (online, 1-21). https://doi.org/10.1177/1741143216659297.

Greany, T. (2018). *Sustainable improvement in multi-school groups* (DfE research report 2017/038). London: Department for Education.

Greany, T., (2019). Approaches to scaling innovations across schools: An analysis of key theories and models. In M.A. Peters & R. Heraud (Eds.), *Encyclopedia of educational innovation*. New York: Springer.

Greany, T., & Higham, R. (2018). *Hierarchy, markets and networks: Analysing the 'self-improving school-led system' agenda in England and the implications for schools*. London: IOE Press.

Gronn, P. (2002). Designer-leadership: The emerging global adoption of preparation standards. *Journal of School Leadership, 10*(1), 552–578.

Hall, G. E. (2013). Evaluating change processes: Assessing extent of implementation (constructs, methods and implications). *Journal of Educational Administration, 51*(3), 264–289.

Hallgarten, J., Hannon, V., & Beresford, T. (2015). *Creative public leadership: How school system leaders can create the conditions for system-wide innovation*. Dubai: WISE.

Hargreaves, D. (2012). *A self-improving school system: Towards maturity*. Nottingham: National College for School Leadership.

Honingh, M., Van Genugten, M., Van Thiel, S., & Blom, R. (2018). Do boards matter? Studying the relation between school boards and educational quality. *Public Policy and Administration, 35*, 65–83. https://doi.org/10.1177/0952076718789739.

Hopkins, D., Stringfield, S., Harris, A., Stoll, L., & Mackay, T. (2014). School and system improvement: A state of the art review. *School Effectiveness and School Improvement, 25*(2), 257–281.

Janssens, F. J., & Ehren, M. C. M. (2016). Toward a model of school inspections in a polycentric system. *Evaluation and Program Planning, 56*, 88–98.

Lubienski, C. (2009). *Do quasi-markets foster innovation in education? A comparative perspective* (OECD Education Working Papers, No. 25). Paris: OECD.

Matthews, P., & Ehren, M. C. M. (2017). Accountability and improvement in self-improving school systems. In P. Earley & T. Greany (Eds.), *School leadership and education system reform*. London: Bloomsbury.

Matthews, P., & Headon, M. (2015). *Multiple gains: An independent evaluation of challenge partners' peer reviews of schools*. London: UCL-Institute of Education, Trentham Press.

Newton, P., & Da Costa, J. (2016). School autonomy and 21st century learning: The Canadian context. *International Journal of Educational Management, 30*(7), 1279–1292.

O'Brien, P. (2015). Performance government: Activating and regulating the self-governing capacities of teachers and school leaders. *Educational Philosophy and Theory, 47*(8), 833–847. https://doi.org/10.1080/00131857.2014.930682.

OECD. (2013). *PISA 2012 results: What makes schools successful? Resources, policies and practices (volume IV)*. Paris: PISA, OECD Publishing. https://doi.org/10.1787/9789264201156-en.

OECD. (2015a). *Education policy outlook 2015: Making reforms happen*. Paris: OECD Publishing.

OECD. (2015b). *Schooling redesigned: Towards innovative learning systems, educational research and innovation*. Paris: OECD Publishing.

OECD. (2018). *Equity in education: Breaking down barriers to social mobility*. Paris: OECD Publishing.

Ozga, J. (2009). Governing education through data in England: From regulation to self-evaluation. *Journal of Education Policy, 24*(2), 149–162.

Perryman, J. (2009). Inspection and the fabrication of professional and performative processes. *Journal of Education Policy, 24*(5), 611–631. https://doi.org/10.1080/02680930903125129.

Sahlberg, P. (2010). Rethinking accountability in a knowledge society. *Journal of Educational Change, 11*(1), 45–61.

Suggett, D. (2014). *Networking as system policy: Balancing vertical and horizontal dimensions*. Paris: OECD CERI.

Waslander, S. (2010). Government, school autonomy, and legitimacy: Why the Dutch government is adopting an unprecedented level of interference with independent schools. *Journal of School Choice, 4*(4), 398–417.

Woods, D., & Macfarlane, R. (2017). What makes a great school in the 21st century? In P. Earley & T. Greany (Eds.), *School leadership and education system reform*. London: Bloomsbury.

Professor Toby Greany is Professor of Education at the University of Nottingham. He was previously Professor of Leadership and Innovation at the UCL Institute of Education, where he was also Director of the London Centre for Leadership in Learning and Vice-Dean: Enterprise. Toby's research is focused on understanding the ways in which educational policy and practice interact and the roles of system governance, leadership agency and evidence in this process. In the past Toby has worked for the National College for School Leadership (NCSL), the Design Council, the Campaign for Learning, the Cabinet Office and as Special Adviser to the Education and Skills Select Committee.

Chapter 5
Case Study of a Cluster in the National Association of Head Teachers' 'Instead' Peer Review in England

David Godfrey and Melanie Ehren

Abstract This chapter examines a case study of a cluster of three primary schools involved in the 'Instead' peer review programme, run by the National Association of Headteachers (NAHT). Interviews of the Headteachers and other school teaching staff were conducted and analysed alongside documentary evidence. We looked at the impact of the peer review network on participating Headteachers, their schools and on other local networks that the schools belonged to. We also examined the interplay between the peer review, the self-evaluation and school inspections. At school level a number of improvements were described to quality assurance, leadership development and other areas. Inspections were found to be a double-edged sword; on the one hand, they motivated schools to engage in peer review, as they would do so to prepare for inspections. On the other hand, inspections motivated a school-based focus instead of supporting schools to develop collective capacity.

5.1 Introduction

Over recent years, the English school system has become increasingly dominated by the themes of school autonomy, strong central accountability and networked improvement. These have been emphasised in White Papers in 2010 (DfE 2010) and that began a rapid increase in so-called Academies (independent public funded schools) and continued in 2016 (DfE 2016). Alongside the autonomy theme, schools have, in this period, been urged to join networks, in particular Multi-Academy Trusts (MATs) as well as others such as Teaching Schools Alliances (TSAs) (Ehren and Godfrey 2017). Even for schools that remained under local authority (LA) control, there has been a strong drive towards networking and partnerships. During this time, peer review has been increasingly used as a form of professional

---

D. Godfrey (✉) · M. Ehren
UCL Institute of Education, London, UK
e-mail: david.godfrey@ucl.ac.uk

© Springer Nature Switzerland AG 2020

D. Godfrey (ed.), *School Peer Review for Educational Improvement and Accountability*, Accountability and Educational Improvement,
https://doi.org/10.1007/978-3-030-48130-8_5

accountability and school improvement (see Chap. 1). While this activity has been sharply increasing over the last decade, the research base is still lagging behind, with the majority of the research on (within school) internal evaluation (Nelson et al. 2015).

This chapter examines a case study of a cluster of primary schools in England involved in the 'Instead' peer review programme, run by the National Association of Head Teachers (NAHT). Interviews of the Headteachers and other school teaching staff were conducted and analysed alongside documentary evidence, including the school's self-evaluations, peer review reports and inspection reports. The research questions we sought to address were:

- What was the impact of the peer review network on the participating Headteachers, their schools and on the other local networks that the schools belonged to?
- What is the interplay between the peer review, the self-evaluation and school inspections?

We examine the external contexts of the schools involved and how these influenced the choice of focus and motivations to participate in this peer review programme. We also briefly outline some relevant theoretical points from the literature in relation to accountability and school evaluation. In our analysis of the case study, we analyse the types of evaluation methodologies employed, how judgements were arrived at and the involvement of users in the evaluation process. Finally, we look at the impact of the review on the schools as a result of their involvement before reflecting on the research questions above.

5.2 Accountability and Improvement Through Peer Review in the English School System

Since 2010 two successive UK governments have set out an agenda based on a school-led, 'self-improving' system. Supported by OECD research, high autonomy coupled with strong accountability has been promoted as the optimal model for system success (OECD 2010, 2011). England is often considered to have one of the world's most autonomous school systems coupled with one of the most high-stakes in terms of external accountability (e.g. Glatter 2012). The extent of this autonomy is the source of some dispute, in that schools' financial freedom to choose service providers for example, may be more evident than schools setting their own curricula or innovations to teaching practice. This can be explained at least in part by pressures to conform in a system of high stakes external accountability (Ehren 2019).

In England, the high-stakes external environment has been described as having deleterious effects on the professional and leadership environment in schools (Gilbert 2012). As Knapp and Feldman (2012) state:

> *A growing body of research documents how educators and students experience external accountability systems. The research makes clear that these system demands can be experienced as onerous, punitive, intrusive, and de-skilling* (p. 668).

The term 'panoptic performativity' describes a constant state of 'inspection readiness' that particularly applies to those schools with short- or no-notice warnings (Perryman et al. 2018 p.147). Simultaneously, many teachers in England belong to a 'post-performative' generation who, while eager to retain a degree of professional autonomy, recognize the need to respond to external accountability measures. Knapp and Feldman define an intersection and interplay between the demands of external accountability and internal accountability (on staff, schools and children) that school leaders in particular need to manage to their advantage. One of the solutions is for school leaders to build strong professional and collaborative responsibility, focusing on learning-centred leadership, and forging strong connections with other school leaders (Knapp and Feldman, p. 674).

However, for strengthened internal accountability to lead to school improvement this requires, "a collaborative culture that combines individual responsibility, collective expectations, and corrective action" (Fullan et al. 2015, p. 4). In order for a school-led system to mature beyond one of compliance to external accountability measures, three key drivers have been proposed: strong joint practice development; partnership competence and collaborative capital (Hargreaves 2012). The first dimension involves joint activity, in which two or more people interact and influence one another, focussing on teachers' professional practice, and developing rather than simply 'transferring' this practice. The second dimension requires a collective moral purpose, combined with high social capital distinguished by trust and reciprocity alongside evaluation and challenge. In combination with these elements, the increase in collaborative capital should enable schools to work together and achieve more than the sum of their parts (Hargreaves 2012).

Seen in this context, collaborative school peer review programmes exemplify a shift in the nature and role of accountability in schools. Earley and Weindling (2004) have proposed four main areas of accountability: 1: pupils parents and the local community (moral accountability); 2: colleagues (professional accountability); 3: employers or government (contractual accountability) and 4: the market, where clients have a choice of institution (market accountability). While all remain important, Gilbert (2012) suggests much greater emphasis needs to be placed now on the moral and professional aspects of accountability and that intelligent accountability based on test and performance data needs to be coupled with an all-important evaluative approach leading to development and school improvement (Gilbert 2012, p. 8–9). Such an emphasis is supported by this comment from the influential OECD Director of Education and Skills, Andreas Schleicher:

> *… devolved decision-making needs to go hand in hand with intelligent accountability. This means moving beyond approaches to external accountability towards building capacity and confidence for professional accountability. Networks of schools can stimulate and spread innovation as well as collaborate to provide curriculum diversity, extended services and professional support* (Schleicher 2013, p. 11).

This kind of moral and professional accountability is akin to the sense of responsibility that professional have for the children in their care (Gilbert 2012 p. 8). This can be more of an affective dimension, more personal, and often may provide an impetus towards formative evaluation, fueled by a motivation to improve policies and practices.

5.3 The English Policy Context

At the time of these case studies, English primary schools were adapting to a number of significant structural, legislative and political initiatives.

In 2010 the White Paper, The Importance of Teaching (DfE 2010) set out the intention of the government to fast track the process by which schools could become academies. These independent state schools, uncoupled from local authorities (LAs), offered the promise of greater autonomy compared to their LA counterparts. At the time of our case studies, forced academisation was not unusual, for schools that were designated to be under achieving, alongside a government push for academised schools to join Multi-Academy Trusts (MATs).

The Ofsted inspection framework had also been recently changed to new ‘light touch’ inspections for schools rated good or above. These would be shorter visits, less frequent and with shorter notice. The presumption was that good schools were likely to remain this way unless new evidence came to light. The intention of these new measures was for schools to spend less time in preparation for visits and to ease the burden of external accountability when schools had earned the right for such autonomy.

A range of new policies was introduced in this period. A recent emphasis on promoting ‘British values’ in schools was also highlighted in the so-called Trojan Horse scandal.[1] Schools were expected to have very clear Spiritual, Moral, Social and Cultural policies (SMSC), to interpret these values for themselves in accordance with the local context. A new assessment process led to the scrapping of standardised pupil progress ‘levels’ leaving schools to decide how to monitor this. A new funding formula was introduced, allocating greater funding to each school per child identified as being from deprived backgrounds (e.g. eligible for free school meals). This so-called ‘pupil premium’ system was intended to re-distribute funding more equitably around the system.

Our interviewees referenced these external policy factors when deciding on the focus for their reviews and in the school development targets.

---

[1] http://www.telegraph.co.uk/education/11244547/Seven-schools-in-latest-Trojan-Horse-scandal.html

5.4 Background to the Peer Review Programme

The Instead peer review model arose initially out of the NAHT's Aspire Partner Schools Programme,[2] piloted in 2013 and evaluated in December 2015 (Neary et al. 2015). According to those involved in developing and running these school partnership programmes, they came in the midst of a climate of dis-satisfaction among Headteachers, peaking in 2012 as Ofsted (Office for Standards in Education) inspections became increasingly high stakes. For instance, Ofsted determined the removal of leadership of a school achieving three consecutive 'satisfactory' grades. Those at the NAHT felt that Ofsted was too adversarial and the inspection system was not felt to lead to improvements. Primary school headteachers had been critical about Ofsted inspections because they had experience of inspectors who had been from secondary or vocational sectors, leading to resentment that they lacked the expertise to make valid judgements about primary schools. The NAHT coordinators we spoke to also felt that some headteachers were not well-equipped to self-assess, and therefore may benefit from having a peer come to validate this process.

The NAHT had funding from the Department for Education (DfE) to develop and pilot the initial model, which was called 'Aspire' with a cluster of schools. In the pilot, schools judged to be 'requiring improvement' by Ofsted 'aspired' to get to a 'good' grade within 3 years. After 2 years of the Aspire programme, which had multiple school improvement strands, 19 out of 30 of the schools that had been inspected during the course of the programme had moved to a Good grade (Neary et al. 2015).

One of the outcomes of the pilot was an understanding that good schools would also benefit from such a programme. The 'Instead' peer review became the solution to this, offering a programme that could lead to school improvements, but where the focus was decided by the headteachers involved, as opposed to pre-determined aims based on Ofsted report recommendations, as was the case for the schools involved in the Aspire clusters. Initially this model was felt to favour schools that were Ofsted rated 'good' or 'outstanding' only, where schools more capable of driving their own improvements could do so through collaboration. The inclusion of lower rated schools was not excluded in the potential development of Instead, however.

The Instead programme was a less-structured programme than Aspire, focusing on smaller clusters of 3 or 4 schools as opposed to 5–10 in Aspire and without the additional network events, coaching, input on school effectiveness and focus on specific pedagogical practices of the latter. The peer review cluster came together primarily to help schools that were good or outstanding and did not involve schools requiring improvement as judged by Ofsted inspections. By January 2015, four (pilot) schools had completed the Instead peer reviews and six were scheduled for a second cluster of peer reviews. There was considerable interest in the NAHT Instead model across the country and more geographical areas were due to be included in the following academic year.

---

[2] http://nahtaspire-co-uk.stackstaging.com/naht-aspire/

5.5 Methodology

The research followed a diachronic case study methodology (Thomas 2011), looking for effects of involvement in the peer review programme over the period of the 2015/16 school year. The effects of the peer review are located within the policy, network and internal school contexts, and accounts from interviews are triangulated with data from documentary sources.

5.5.1 The Case Schools and Staff

NAHT coordinating staff were approached for permission to take part in the study in January 2015 and they agreed to broker contact with a cluster due to come together that year.

The NAHT agreed to waive their anonymity, although have clarified that they no longer coordinate these peer reviews, which have evolved into a new iteration of the original Aspire programme.[3] Details of the three schools are given below in Table 5.1; school, staff and names of the English counties from which the schools came, have been pseudonymised.

Our three case study schools completed their reviews between September 2015 and July 2016. The headteachers were all members of the NAHT and two of them had met previously. The NAHT assisted in the formation of the cluster but the three Headteachers had provided information to the NAHT staff about what sort of schools they were looking to work with. One of the cluster schools was based in a large city, one in a large town and the other in a more isolated, rural setting. They were all primary schools and all had been previously graded ‘Good’ (two in 2013 and one in 2014). They were located within the same general part of England although none were in the same local authority. They had in common at least a

Table 5.1 Summary details of case schools involved in the NAHT peer review cluster

School name	Headteacher and date of appointment	Local authority area	Setting	Number of pupils	Most recent Ofsted grade and date	Date of instead review visit
Holy primary	Alice (2008)	Countryshire	Rural	650	Good (Oct 2013)	2nd Feb 2016
Roundtown primary	Samantha (1999)	Metropol	Urban	237	Good (Jan 2014)	15th March 2016
Greenleigh primary	Evelyn (2004)	Landshire	Urban	440	Good (Jan 2014)	24th May 2016

---

[3] https://www.naht.org.uk/membership/special-partner-offers-for-members/school-improvement-programmes/naht-aspire/aspire-peer-review/

relatively high level of pupils for whom English was an additional language (EAL) and/or pupils on free school meals (FSM).

The lead reviewer (LR) (known here as 'Ross') was chosen after the formation of the three primary school headteachers into the cluster and this was decided by the headteachers themselves.

Details of the instead peer review process itself can be seen in Box 5.1 below:

Box 5.1: The Instead Peer Review Model

The Instead matrix consists of four focal areas: Learning and Teaching, Pupils, Community, and Leadership. These are reviewed in a matrix against: The quality of school vision and strategy, the quality of school analysis, and the quality of school delivery.

Schools work in clusters of 3 or 4 and review each other in turn one in school year. Most schools in Instead come from pre-existing clusters ranging in level of permanency and formality. Some schools were put together into a cluster, helped by a coordinator at NAHT, as is the case for the cluster studied for this project. No grades are given in the review rather, the framework and report addresses: 'actions needed and priorities' and 'what should be prioritised, developed, maximised and sustained?'. Lead reviewers (LRs) join the other members (one from each school in the cluster) for 2 day visits for the review at each school.

LRs are taken from outside the cluster, selected from a pool (of around 10 LRs in early 2016). Headteachers of participant schools attended a recruitment meeting and there was an additional training day for those who decided to take part as LRs. All LRs have some prior experience of inspection or peer review.

Once the cluster of schools is formed, the cluster visit schedules are decided between the headteachers and the lead reviewer. Clusters are required to send review dates to the NAHT and then they are reminded to make a 'scheduling request' to Ofsted to avoid their Instead review date. Prior to each visit, the school to be reviewed completes the self-evaluation matrix, addressing 'where they are at' for each part of the grid, adding evidence and performance data that they send to the LR. They also decide on the focus of the visit; e.g. to look at the quality of provision for students for whom English is an Additional Language or to look at the assessment policy across the school. Each school in the cluster takes turns being reviewed. Each review lasts for two whole days and on the third day the report is brought together by the LR in communication with the host school headteacher.

The school has a right to reply after the reviewers have completed their report and given their findings. There is also a section in the report for the school to say what they have learned and to give their reactions to the review and how they felt about it. The reviewers can add their evidence in a different colour to the school if theirs is in disagreement. Although reviewers do observe lessons, they do not grade them, and usually they would do these in tandem with a member of the school leadership team.

Table 5.2 Data collected with dates

School	Interviews (dates before and after peer review visit)	Other data
Holy primary	Jan 2016/June 2016	Peer review report
	Head teacher	Ofsted report
	Deputy head	DfE data
	Two assistant head teachers	
Roundtown primary	Jan 2016/June 2016	Peer review report
	Head teacher	Ofsted report
	Deputy head	DfE data
	Year 5 teacher	Email communications
	Pastoral achievement coordinator	
Greenleigh primary	Jan 2016/Nov 2016	Peer review report
	Head teacher	Ofsted report
	Deputy head	DfE data
	Phase leader for years 3 and 4 and English	
	Phase leader for years 5 and 6 and Maths	
Lead reviewer	Jan 2016 /July 2016	

Thirteen people were interviewed from the cluster, each one in January of the year before the review visit and then approximately 2–6 months after the review. Of the 13 interviewees, these included the Head teacher of each of the three primary schools and the lead reviewer for the cluster. We also asked to see other staff significant to each school review, usually other senior leaders or staff with areas that the peer review would particularly focus on. Other data analysed also included the three Instead review reports, records of email communication between members of the cluster and other publically available data such as Ofsted reports and DfE data. See Table 5.2 above for a summary of data collected. We intended to capture evidence about the effect of the peer review in relation to an Ofsted inspection but none occurred within the time period of this project.

5.5.2 *Data Analysis*

The variables examined in this case study were adapted from the wider European Union comparative study.[4] These looked at ***contextual factors*** such as the ***external (network) context***, including the socio-economic context of the network as a whole, the legislative context and the structure of the national school system and policies that were relevant to the primary schools involved in the case study. We also looked at the ***school context,*** including data on the nature and size of the school's intake, staff turnover, the structure of the leadership team, the socio-economic local context, other specific challenges and the extent of local authority support. Finally we looked at ***other networks***, alliances or partnerships that each school was also involved in, including other peer review arrangements.

---

[4] https://www.ucl.ac.uk/ioe/departments-and-centres/centres/centre-educational-evaluation-and-accountab/research/inspections

Guided loosely by Provan and Kenis's (2008), description of characteristics of networks, data were also analysed in terms of the ***structural*** aspects of the network, i.e. about governance, size and geographical spread; the ***relational*** aspects, i.e. levels of trust, formality of relationships and perceptions about hierarchy between the members of the network; and also ***collaboration***, i.e. frequency of communication and types of information sharing. The extent to which these three aspects affected the formation of the network and influenced its functioning were analysed. In addition, we looked at the motivation for joining a peer review cluster and the focus of their reviews.

We were also interested in the ***evaluation practices*** that participants reported before, during and after the review visits and these were sub-divided into classifications about ***methodology, valuing and judging and user involvement***. *Methodology* referred to the specific evaluation practices used in the conduct of the review, before, during and after each school review visit, including the type of data collected and analysed. *Valuing and judging* concerned the ways that judgements were arrived at and how these were received and valued by the school personnel. *User involvement* looked at which stakeholders were involved in which phase of the evaluation and to what extent, including the definition of the focus and scope of the review; involvement in the self-evaluation, the in the review visit itself and in reaching judgements and in the final review report.

Finally we looked at outcomes. These were initially divided for coding into, ***network-level outcomes***, ***school-level outcomes*** and ***dysfunctional effects.*** For the former, we coded for the production of knowledge to solve problems relevant for the entire network and that went beyond the remit of each individual school. Such examples could include: sharing resources, joint Continuing Professional Development (CPD), improvement activities, support and joint initial teacher training. However, this code was later expanded to also include the effects of the peer review experience on the school's existing networks. School-level outcomes include follow-ups to the recommendations of the review and any other outcomes reported before or during the visit or as a result of taking part in the review. Dysfunctional (network) effects could include shifting dysfunctional teachers across the network, increasing the salary of network managers, making profit or increased competition between schools.

Two researchers on the project coded one interview transcript using these variables and inconsistencies were addressed to improve the internal reliability of the coding. Qualitative data analysis software, NVivo version 11 was used to code and analyse the data.

In the first stage of coding only the above themes were coded and the researcher made numerous coding notes to inform sub-categories (child nodes). As a result, the theme 'evaluation and inspection practices' was further sub-divided into numerous subordinate nodes as further detail emerged in the analysis. Notable among these were comparison with other reviews or with Ofsted inspections. Ways in which staff made judgements (valuing and judging) were also further divided into 27 subordinate nodes, including comments about the subjective nature of judgements and comparisons made by reviewers to their own school.

School level outcomes, reported below, come from within the themes of 'evaluation and inspection practices' and 'relationships, collaboration, structure of the network'. Results at the level of student attainment were not taken into account.

5.6 Motivation and Focus of Reviews

All three Headteachers mentioned the impact of the academisation agenda to their school and how this impacted on their decision to join this peer review scheme. They bemoaned the potential loss of the strong links to the community that had been cultivated over many years at their schools and also the fear of being forced to become an academy and/or part of a multi-academy chain. Samantha raised the issue that the government had proposed introducing a new category of 'coasting schools', and were using this as a tool to force many more to become academies by virtue of underperformance. In Landshire, Evelyn had heard of cases from local schools where forced academisation had proved a painful and divisive process.

> *The head was basically frogmarched out with no warning one afternoon and not allowed to take anything with her, and nobody had any information, it really was badly handled. So everybody's heard the horror stories (Evelyn, Greenleigh Primary).*

Finding appropriate local collaboration models for schools was a way for the school to choose who they wanted to work with and under what terms. The consequence of the squeeze on LA finances was that schools needed to buy in more support themselves, increasingly scouring the market for what was most appropriate to their needs. By remaining under LA control, these schools were at risk of becoming increasingly marginalised while also being scrutinised more heavily by the LA and asked to take on an ever-larger burden of local issues, such as providing support for struggling (LA) schools. Given the lack of LA resources, the NAHT peer review was seen as a favourable alternative to other peer review or school improvement collaborations, some other local models being described as very 'Ofsted led' or less developmental in nature.

The headteachers were mindful that Ofsted inspectors on the new 'light touch' inspections (for schools rated good or above) would have little time to come to the school or to discuss the context surrounding the school data. These short notice inspections now meant that schools would not have the preparation time for inspection visits either. As a result, senior leaders at these schools felt the need to be 'ever-ready' to have the kind of discussions about data and progress of their students, that would be required in discussion with inspectors. The peer review was anticipated as a way to rehearse these conversations and seen as integral to new senior leaders' professional development.

For Samantha at Roundtown Primary school, the review was also seen as part of her succession planning, knowing that she was going to retire at the end of the year and for both her deputy and other (especially new) staff needed to have ownership of the priorities identified in the Instead review. At Greenleigh Primary, Evelyn was conscious that some new members of the leadership team would benefit from the practice of being 'held accountable' for their areas of responsibility. The practice of explaining the data underlying individuals' areas of responsibilities was seen as a key part of leadership development.

The national reforms and policy changes mentioned above influenced the focus of the reviews themselves, these included the need for schools to promote '*British values', changes to assessment in primary schools and the new funding formula for schools.* Holy Primary presumed that Ofsted inspectors would want to see very clear displays around the school that outlined their SMSC policy in explicit terms and hoped that the Instead peer reviewers would be able to give them feedback on their attempts to meet these requirements. As a consequence of the assessment changes, schools were now looking for new ways to measure and track progress, including the use of computerised management information systems. The schools varied in their adoption of particular methods and were keen to have feedback on the relative pros and cons of these from their peers. Samantha, at Roundtown Primary was keen for the reviewers to evaluate how they tracked progress in order to help them improve their pupils' learning. The new funding formula meant that Roundtown Primary and Metropol School (the LR's school) would see their funding cut while Alice and Evelyn would potentially see greater funding in this re-allocation process. Samantha felt that hers and Ross's schools would be 'unsustainable' with these cuts, given their current dependence on pupil premium funding. Therefore, discussions about the use of pupil premium funding were important for the peer review visits and these schools also expected Ofsted to be interested in finding out in detail how this money was being used.

5.7 Relationships, Collaboration and Structure of the Network

The informal nature of the cluster, the credibility of all members within it and the high levels of trust, were all seen as crucial aspects to the functioning of this peer review network.

The cluster came together informally as the headteachers were all members of the NAHT and had found out about the new peer review model that was been offered to primary schools. The three schools were geographically within the same general part of England although none were in the same local authority.

Although there were some important structural differences between each school, in terms of size, rural or urban, ethnicity and funding, they also shared similarities such as having a lot of deprived students and many with English as an Additional

Language (EAL). The distance they each had to travel to the other schools was not seen as a big problem during the year, as long as the collaboration was seen to be (potentially) useful. However, Ross did acknowledge that the gap of ‘150 miles’ with his school, meant that future collaboration, apart from by email, always seemed unlikely. Nevertheless, a view shared among all the Heads, was that there was significant value in collaborating with schools who were not existing or natural partners, as they would come to the review without preconceptions and with a fresh perspective. The ability to confront shared challenges in an open relationship of trust was seen as lacking in some of the schools’ local partners too.

Across the three reviewed schools and also with Ross the lead reviewer, there was a high level of trust and mutual respect; Ross described their ‘over 50 years’ of shared experience.

This had been helped as the head teachers of the schools had got to know each other in two earlier meetings before the first review. They shared a similar agenda and values and these were key areas of glue in the collaboration; in particular, the best interests of the children ultimately drove the review. The open discussion and dialogue enabled the Head teachers to freely ask advice of the other Heads; there was no perception of a hierarchy, competition or dominance by one single contributor to the review.

5.8 Evaluation Practices in the Reviews

Prior to the visit**,** each headteacher sent out documents for the team to review, along with their own self-evaluation. These included: current school development plans; the latest Ofsted inspection report; the latest school self-evaluation; RaiseOnline[5] data and Ofsted ‘Dashboard’[6] data. During the visits the reviewers looked at pupils’ work; conducted lesson observations or ‘learning walks’ and interviewed school leaders, teachers, teaching assistants, other staff, governors, parents, pupils (including the school council) and looked at the school environment (e.g. wall displays) and in one case, attended a school assembly. In this sense, the evidence gathered was very much in line with what might be expected in an external inspection.

At the end of each review visit Ross (the LR) was responsible for collating the evidence and he would then send the draft report out to all three headteachers to check that they were in agreement. A section of the report allowed the host school to describe what they had learned from the visit and also to comment on the reaction of wider staff about the usefulness of the review days. The right hand column of the report had a summary of the school, the data collected and the key findings in relation to the validation of the self-evaluation matrix. The second page outlined the key strengths of the school in relation to learning and teaching, pupils, community and leadership. The last page then gave short bullet-pointed recommendations to the school in terms of what it should prioritise, develop, maximise and sustain. The reports were quite short, at around 800–900 words in total.

---

[5] http://www.raiseonline.org. School performance data now being replaced by ‘Analyse School Performance’.

[6] https://www.gov.uk/government/collections/using-ofsteds-inspection-dashboard. Inspection data now being replaced by a new Inspection Data Summary Report.

5.9 Involvement of 'Users'

The extent to which each school involved the wider staff in the self-evaluation and review day varied across the schools. In terms of the focus of the review, the headteacher and the Senior Leadership Teams (SLT) largely decided the key areas they wanted the reviewers to look at. At Holy Primary School, Alice asked each of the senior leaders to complete the NAHT self-evaluation matrix and then the team came together to try to reach consensus and decide on what the focus of the review should be. At Greenleigh Primary they conducted the self-review initially among senior leaders, and then in an In-Service Training (INSET) day, asked departments to conduct and submit their responses anonymously, giving them a sense of where there were discrepancies. The aspects of the Instead self-evaluation grid led to thinking about areas not previously covered in the schools' normal self-evaluations, such as those on 'whole school community' and 'the quality of school vision'. Therefore, the process of discussion was seen as useful in raising awareness as much as it was about deciding the focus of the review. The host headteachers put together the timetables for the visits, and others helped in terms of setting up meetings with staff on the first morning of a visit and how to help the visitors find their way round. However, there were variations in the extent to which other staff were aware of the other reviewers' backgrounds or the exact purpose of the visit.

During the visit, members of the SLT at the host school would walk around with the reviewers and attend some of the lesson observations/learning walks with the visiting reviewers. This was seen as a way to validate each other's judgements, and importantly, that of the host school SLT. The review visit final feedback session at each school included the SLT of the host school; other staff were informed about outcomes after the visit when SLT shared the report with them. The report itself had a section for the host school to respond to the outcomes of the review and reflect on their learning from the day. Reports went out also to governors and in at least two of the schools, the challenge and support provided by the review team to the governing body was seen to be a particular area of strength. Reports were not shared with parents; Alice at Holy Primary felt that without the right context and/or written in the correct language, the report could give a misleading impression.

5.10 Valuing and Judging in the Review Visits

At the end of each review visit, the Lead reviewer summarised the findings of the review team in a short report; as well as a summary of the school and of its strengths and weaknesses, there were recommendations that identified areas to prioritise, develop and maximise. Dialogue between the school senior leaders and the visiting reviewers about the judgements were reached as they went through the day, in the final meeting on day two, and in the formulation of the final report. Schools sought validation of their self-evaluation and this included endorsing the good points of the school, enabling these to be 'celebrated'. In addition to the validation of the self-evaluation, the reviewers gave feedback on the school development plans, based on the evidence they had viewed. A summary of recommendations for each school is show below in Table 5.3.

The production of the end report invited a degree of negotiation, reflecting a relationship of equals. There was a section for the school to comment on the review and also the host headteacher was able to contest the content of the report if they wished to. In the case of Roundtown Primary, there was some disagreement between Samantha and Ross the LR about the content of her school's draft report with respect to whether part of their curriculum was able to differentiate for higher and lower achieving pupils. These disagreements were shown in a series of email exchanges and a compromise was finally reached in how the report was worded. Such an exchange proved a clear contrast to how external inspection reports are put together, where host schools are only allowed to challenge factual details.

It is also important to note that learning was gained as much from visiting and reviewing schools as it was through the hosting of a visit. All the headteachers expressed their impression of this as a highly personally and professionally developmental process.

The interviews from the participating schools suggested that reviews were 'non-judgemental'. However, the use of subjective comparisons and anecdotal evidence by reviewers was sometimes valued and other times not. Senior leaders at Holy Primary and the Headteacher at Roundtown Primary criticised the visiting Heads for making references to their own school that were either seen to ignore differences in the context of the host school:

> *early years provision looked different to theirs, and we know early years is an area we've been working on, but you can't keep saying oh at my school we've got this. You can't do that, you've got to look objectively and not compare it to your own school* (Mariana, Assistant Headteacher, Holy Primary).

By contrast, Shaun, the Deputy Headteacher at Greenleigh noted ways in which subjective comparisons were sometimes welcomed:

> *It seems to me they* [the reviewers] *were doing something that an Ofsted inspector wouldn't do, which was they were comparing our school with their own schools…. Whereas a serving head teacher can go in and say I'd be more than happy for my key stage two classrooms to look like this and feel like this, and how do you manage that?*

Table 5.3 Summary of recommendations from each peer review visit report

School	Recommendations
Holy primary	**The school should prioritise:**
	Early years provision, including agreeing the underpinning philosophy, and a review of the curriculum
	Analysis of the impact of interventions in place for SEN pupils
	The school should develop:
	The commitment of the relatively new governing body to ensure they understand their areas of oversight
	The school should maximise:
	A review of the homework policy to clarify its purpose
	Pupil progress meetings so that teacher's have ownership of the process
	Support and scaffold for maths to ensure a number rich environment
	Marking and feedback in maths
Roundtown primary	**The school should prioritise:**
	The marking and feedback policy to ensure consistency across the school.
	Access to, and ownership of the data for all relevant pupils by all staff
	The school should develop:
	A new leadership structure that is more streamlined and includes a named person responsible for the use of pupil premium funding
	The school should maximise:
	The impact of the leadership of English and Maths, making clear the expectations for monitoring and evaluation.
	Differentiataton of the curriculum to meet the needs of all pupils.
Greenleigh primary	**The school should prioritise:**
	The marking and feedback policy to ensure that all teacher marking is effective, including how pupils respond to questioning
	The school should develop:
	A review of staffing and spending for the pastoral team to ensure that it can continue to provide excellent support for vulnerable pupils
	The relatively new leadership team including plans to continue their own professional development and a focus on coaching and mentoring.
	The school should maximise:
	Overall attendance as well as ensuring that it has an awareness of the context and needs of specific groups of pupils and families

The absence of a final public report (and no overall 'grade') allowed school leaders to be honest about their school's weaknesses, inviting advice on how to improve; accounting for the success of the school as well as its areas for development and opening up collaborative and supportive dialogue between members of the reviewing team.

At Holy Primary School, where they had experienced an alternative peer review in their local cluster a few days later, interviewed staff were overall less positive about their experiences with the review visit which they described as woolly and lacking clear action points:

> *I was hoping for more of a this is where you're at, and these are your development points, these are the areas that you, so that you can actually then build upon that, action plans etc, and if I compare that with the Triad [their local peer review cluster] we've just had that's what I've come out with, very clear, structured, development points* (Charlotte, Holy Primary).

Overall, especially at Holy Primary, there was a sense that the review 'merely' confirmed what the school already knew. While reassuring, they also felt that they had not been sufficiently challenged during interviews by the visiting Heads. There was a sense that many staff had built themselves up for a process that would 'put them on their mettle' in the way that an Ofsted inspection team would and when this did not prove to be the case, the review had been anti-climactic:

> *So in our heads that's what we are ready for, and prepared for, and when it doesn't turn out to maybe be like that then you are maybe left thinking oh that's a bit…wishy washy. But is it because you're thinking it's going to be an Ofsted, a pre-Ofsted, and when it's not, when it is more of a peer review, which is what it's called, but you try and tell teachers when you've got headteachers coming in and you are walking around watching them it's not an Ofsted, it's a very hard thing to get across. Very hard* (Mariana, Assistant Headteacher, Holy Primary).

Despite the clear separation of purpose and methodology between 'Instead' reviews and 'Ofsted' inspections, the experience and training of all the headteachers in Ofsted methodology influenced the evaluation practices during the review visits. Shaun at Greenleigh Primary, confirmed that this crept into judgements that the reviewers made during their visit to the school:

> *… if I was an Ofsted inspector the school would be at least good with outstanding features, and you are so far from outstanding, and these are the things you would need to do. So they were talking in those terms.*

There seemed an ambivalence about such comments too. Clearly they were made with god intentions and partly seen to help the school prepare for an inspection visit. On the other hand, an Ofsted style judgement may have detracted from the wider ranging and more developmental aims of the peer review.

5.11 The Impact of the Reviews on Schools and Participants

We interviewed the headteachers and other staff of each school 2–6 months after the review about follow-up actions and the impact of the peer review. Table 5.4 below summarises their responses.

There were some clear patterns in impact and these reflected a conscientious follow up on the recommendations of the reviewing team. For instance, both Greenleigh and Roundtown Primary Schools made changes to leadership structures and leadership development for new members of the leadership team. However, it was difficult to distinguish between actions that schools had initiated due to the review and those that were already planned but that had been validated by the review. At Greenleigh

Table 5.4 Impact of the peer reviews as reported in staff interviews 2–3 months after each review

School	Impact or action taken
Greenleigh primary	Sponsoring two new leaders to take national professional leadership qualifications and providing in-house mentoring in their new roles
	Greater school preparedness for an Ofsted inspection, especially of the two new SLT
	Changes to the pastoral team
	Changes to marking system to ensure it was better implemented and understood.
Roundtown primary	Following up suggestions to work on the assessment policy to get in place for the new school year
	Following up recommendations about pupil premium and monitoring its impact, including keeping data (case studies) to evidence strategies employed with vulnerable students
	Changes to leadership structure
	Change in practices at middle leadership level regarding holding meetings for year groups
	Improving the use of data systems by all staff
	Evidencing how the curriculum differentiates for learners at each end of the achievement scale
Holy primary	Governors were more aware of their accountability responsibilities for their areas in ways that Ofsted could question them
	Validated school's already existing commitments
	Confirmed school's determination to use data such as test results to do with early years, phonics, and key stages one and key stage two to steer school development plans

and Roundtown schools, the reviews were seen to be useful in strengthening resolve, and providing further triangulation for the schools' self-evaluations:

> *It told us what we already knew to a certain extent, but it gave us a clear view on what was needed rather than, we kind of knew that marking, when we, through book trawls and things like that, there were different people doing different things, and we'd spoken so many times on we need to do this, we need to do that, then obviously when it comes up as a target for development specifically we said right, we need to get it sorted and get it done properly. It gave us I think a clear vision and clear focus and little shake to get things going. Excellent (*Carl, at Roundtown School).

At Holy Primary School staff suggested most strongly that they had not taken actions in response to the review but would rely on other data:

> *There was initial verbal feedback in the staff meeting, which was around the strengths mainly, and then there was a full report, then that was shared, which has development points, but I wouldn't say it was something that the school has focussed on since (*Charlotte, the Assistant Headteacher at Holy Primary).

This contrasting view by Holy Primary participants appears to reflect both lack of trust in the reviewing skills and judgement of the visiting headteachers and the lack of rigour and challenge they perceived from the process. Partly this may also be explained by being the first school reviewed and that they had been insufficiently briefed staff about what to expect.

5.12 Conclusions and Discussion

Our two central research questions were: What was the impact of the peer review network on the participating headteachers, their schools and on the other local networks that the schools belonged to? And; what is the interplay between the peer review, the self-evaluation and school inspections?

To deal with the first question, we found a number of benefits to the schools in terms of improvements to quality assurance, leadership development, greater awareness of accountability roles, how to deploy funding effectively and on the monitoring of student progress. School staff also felt more prepared for a future Ofsted inspection, albeit Holy Primary staff felt reviewers had been too soft on them. The peer reviews were generally not seen to be disruptive to school life, although there was a cost to time of the headteachers involved and in terms of stress levels to some of the participating schools' staff.

We found little or no network level outcomes. Those that we did find occurred largely before the reviews took place and included some exchange of resources and ideas to do with assessment, the use of pupil premium funds and how to track pupil's progress. One exception was the visit of Maths and English leaders from Greenleigh School to Holy Primary School during the same academic year to learn about good practice in these areas.

This research raises questions about whether peer review can build genuine lateral accountability and professional collaboration when incentives and high stakes are organized around individual school performance. As we have seen elsewhere with formal networks (Ehren and Godfrey 2017), perverse effects occur when such networks are held to account by the units within them rather than the impact they achieve as a connected whole. In this cluster, the motivations were at the individual school level and the sense of collective responsibility to the pupils was more 'philosophical' than amounting to concrete commitments to collaborate in the future. Even in the English system, where schools are incentivized to collaborate, our case study schools did not continue their peer review work. The network had been set up for the purpose of a one-year cycle of reviews and the lack of geographical proximity meant the apparent end to collaboration after the year. Provan and Kenis's (2008) work on conditions of effective networks would suggest that not only was distance a problem, but also that the overall governance structure was too loose, particularly in a system that is increasingly formalizing such structures.

However, a by-product of the schools' involvement in the peer review was the promotion of this model for working in other nascent or existing local partnerships. Involvement in the peer review also inspired one school to look further outwards from their local authority area, to link up with another school in relation to improving early years' provision. Therefore, it may be that in the long term, the school found these other networks and alliances productive in supporting longer term improvements having used the peer review as a learning experience to build on and plan future actions.

However, it was difficult to say how much the peer review had motivated any particular changes or merely validated the school's existing development plans. The peer review provided additional 'triangulation' of these plans while other data was likely to be very influential too. It is unclear how much the peer review was likely to have a longer term impact on the school or whether it would lead to (or contribute to) an improved Ofsted Inspection grade.

In terms of the benefits of the peer review to individuals, these included the challenge and support it provided to senior and middle leaders and the governors. A major part of this was being held account for their areas of responsibility and in some cases staff valued the coaching role that some of the reviewers took with them. This accords with the notions of accountability as a social practice in which, "One has to answer questions about what has happened within one's area of responsibility and provide a story or an account of practice; what has happened and why it has taken place" (Møller 2009, p.39). Such narratives are often lost in short external inspection visits where inspectors are driven largely by data and less inclined to listen to reasons underlying the actions taken by school leaders and teachers.

To address the interplay of external evaluations and peer review; inspections are a double-edged sword. On the one hand, they motivated schools to engage in peer review as they would do so to prepare for inspections and as their engagement in peer reviews would assess the leadership of the school (one of the inspection standards). On the other hand, inspection restricted peer reviews in that the focus of was on the inspection criteria instead of a wider understanding of school quality or specific school policies. Inspections also motivated a school-based focus instead of supporting schools in developing some level of collective capacity and benefit from being part of a network (e.g. in generating the type of network-level outcomes we studied).

While anxiety about future Ofsted inspections provided a backdrop to the Instead review, there was an acceptance by most of the participants that inspections had an (important) place in the education system, and that being prepared for them was a key skill for a school leader. Indeed, two of the Heads were Ofsted trained, and one went to work for Ofsted full time as an HMI at the end of the academic year, the latter reflecting the inspectorate's drive since 2014 to recruit more experienced senior leaders with recent practice experience.[7] The very centrality of Ofsted to professional life in schools was clearly internalised by the Heads and senior school leaders, leading to a pervasive influence in the Instead preparation, process and follow up. Evelyn's point below, sums this up well:

> *… we didn't really talk about Ofsted descriptors, or Ofsted categories. Although if you get four headteachers together I don't think it's long before Ofsted gets mentioned.*

---

[7] https://montrose42.wordpress.com/ofsted-recruiting-more-serving-headteachers-to-inspection-roles/

The ambivalence towards peer reviewers adopting this inspection-rehearsal stance seems redolent of the reaction we might expect from the post-performative generation (Wilkins 2011), mentioned earlier in this chapter. One view was that the evidence gained and judgements arrived at from the review could be used as potential leverage if an Ofsted inspector wished to question particular aspects of the school covered by the review. This represents a tactical uses of peer review evidence in contrast to the formative intentions of the Instead programme.

The role for peer review programmes in the development of school leaders was particularly strongly voiced in these case studies. Further research could usefully look at the peer reviews in more structured, permanent and formal school clusters, such as MATs in the English context. These kinds of more formal networks may be able to use peer evaluations as an integral part of a collaborative improvement process fuelled by a more concrete sense of collective responsibility leading to greater network level impact.

References

DfE. (2010). *The importance of teaching* (Report number: 7980). London: Department of Education.

DfE. (2016). *Educational excellence everywhere* (Cm 9230). London: Department of Education.

Earley, P., & Weindling, D. (2004). *Understanding school leadership*. London: Sage.

Ehren, M. C. (2019). Accountability structures that support school self-evaluation, enquiry and learning. Ch. 3. In D. Godfrey & C. Brown (Eds.), *An ecosystem for research-engaged schools: Reforming education through research* (pp. 41–56). London: Routledge.

Ehren, M. C., & Godfrey, D. (2017). External accountability of collaborative arrangements; a case study of a multi academy trust in England. *Educational Assessment, Evaluation and Accountability, 29*(4), 339–362.

Fullan, M., Rincón-Gallardo, S., & Hargreaves, A. (2015). Professional capital as accountability. *Education Policy Analysis Archives, 23*(15), 1–17.

Gilbert, C. (2012). *Towards a self-improving system: The role of school accountability*. National College for School Leadership. Accessed online 23rd February 2019. http://www.learnersfirst.net/private/wp-content/uploads/Think-Piece-Towards-a-self-improving-system-the-role-of-school-accountability-Gilbert.pdf

Glatter, R. (2012). Persistent preoccupations: The rise and rise of school autonomy and accountability in England. *Educational Management Aadministration & Leadership, 40*(5), 559–575.

Hargreaves, D. H. (2012). *A self-improving school system: Towards maturity*. National College for School Leadership. Accessed online on 25th February 2019 at: https://dera.ioe.ac.uk/15804/1/a-self-improving-school-system-towards-maturity.pdf

Knapp, M. S., & Feldman, S. B. (2012). Managing the intersection of internal and external accountability: Challenge for urban school leadership in the United States. *Journal of Educational Administration, 50*(5), 666–694.

Møller, J. (2009). School leadership in an age of accountability: Tensions between managerial and professional accountability. *Journal of Educational Change, 10*(1), 37–46.

Neary, S., Dodd, V., & Radford, R. (2015). *NAHT Aspire pilot evaluation* (Final Report). Derby., Institute of Education. University of Derby.

Nelson, R., Ehren, M., & Godfrey, D. (2015). *Literature review on internal evaluation*. London: Institute of Education.

OECD. (2010). *PISA 2009 results: What makes a school successful? Resources, policies and practices* (Vol. IV). Paris: Organisation for Economic Co-operation and Development.

OECD. (2011). *PISA in focus*. Paris: Organisation for Economic Co-operation and Development.

Perryman, J., Maguire, M., Braun, A., & Ball, S. (2018). Surveillance, governmentality and moving the goalposts: The influence of Ofsted on the work of schools in a post-panoptic era. *British Journal of Educational Studies, 66*(2), 145–163.

Provan, K. G., & Kenis, P. (2008). Modes of network governance: Structure, management, and effectiveness. *Journal of Public Administration Research and Theory, 18*(2), 229–252.

Schleicher, A. (2013). Is the sky the limit to educational improvement? In F. O'Toole (Ed.), *Why education matters: The importance of education to Ireland's economy and society* (pp. 9–14). Accessed online (11th June) at: http://www.tara.tcd.ie/bitstream/handle/2262/67338/jacb201327.pdf?sequence=1#page=10.

Thomas, G. (2011). A typology for the case study in social science following a review of definition, discourse, and structure. *Qualitative Inquiry, 17*(6), 511–521.

Wilkins, C. (2011). Professionalism and the post-performative teacher: New teachers reflect on autonomy and accountability in the English school system. *Professional Development in Education, 37*(3), 389–409.

Dr. David Godfrey is an Associate Professor in Education, Leadership and Management at UCL Institute of Education in London and the programme leader for the MA Educational Leadership. He was co-director of the Centre for Educational Evaluation and Accountability until 2018 and was a lead inspector for the Independent Schools Inspectorate. An advocate of research-informed practice in education, his projects and publications include research-engaged schools, school peer review, inspection systems and lesson study. In July 2017, David was acknowledged in the Oxford Review of Education as one of the best new educational researchers in the UK.

Melanie Ehren is a Professor in the Governance of Schools and Education systems, director of research institute LEARN! and honorary professor at University College, Institute of Education. Her academic work focuses on the effectiveness of accountability and evaluation systems and aims to contribute to a greater understanding of the interplay between accountability and the broader education system in tackling inequality and improving student outcomes. Melanie has led a number of international research council and EU funded projects in the areas of school accountability, inspection and evaluation. This includes an ESRC funded study on 'Accountability, capacity and trust to improve learning outcomes in South Africa'.

Part IV

Peer Review in Unfamiliar National Contexts: Successes and Challenges

Chapter 6
Peer Review Network of Schools – Lessons from Innovative Practice in Bulgaria

Rossitsa Simeonova and Yonka Parvanova

Abstract This chapter presents an innovative model and practice from Bulgaria – a peer review of a network of schools in Sofia city; we reflect on case study results, and the effects and benefits identified for all stakeholders involved. The network was established as a voluntary entity to exchange good practices and provide mutual support for school improvement. The peer review process was based on a self-evaluation and peer review thematic framework developed by the schools in the network. The peer review process was followed by thematic inspection done by the local inspectorate in order to verify and legitimise the self-evaluation and peer review findings, and the school and network improvements. The inspectorate adapted the self- and peer evaluation criteria developed by the network for the purposes of the inspection procedure and thus providing more valuable recommendations in tune with network's agenda for improvement of parental involvement as one of the recognized priorities of the network to be enhanced. The case study results showed that the implemented peer review has improved principals' and teachers' evaluation competencies and they were highly satisfied by their experience. School principals and members of evaluation teams stressed out on the valuable opportunity to compare their achievements with other schools, their efforts to be recognized by other colleagues and to cooperate to improve. They recognized the peer-evaluation as a collaborative and worthy learning process producing joint evaluation products and know-how with potential to be disseminated and adapted by other schools and networks as innovative model and beneficial practice for all involved.

6.1 Introduction

This chapter presents an innovative model and practice from Bulgaria – a peer review of a network of schools in Sofia city; we reflect on case study results, and the effects and benefits identified for all stakeholders involved. The network was established as

---

R. Simeonova (✉) · Y. Parvanova
Sofia University "St. Kliment Ohridski", Sofia, Bulgaria
e-mail: r.simeonova@fp.uni-sofia.bg; y.parvanova@fp.uni-sofia.bg

© Springer Nature Switzerland AG 2020

D. Godfrey (ed.), *School Peer Review for Educational Improvement and Accountability*, Accountability and Educational Improvement,
https://doi.org/10.1007/978-3-030-48130-8_6

a voluntary entity to exchange good practices and provide mutual support for school improvement. The peer review process was based on a self-evaluation and peer review thematic framework developed by the schools in the network. This was followed by a thematic inspection by the local inspectorate to verify and legitimise the self-evaluation and peer review findings, and the school and network improvements.

This innovative practice was implemented as part of an Erasmus+ Key Action 2 Strategic partnerships' funded project titled 'Polycentric inspections of networks of schools' (09.2014–08.2017) accomplished in partnership with four research teams from England, the Netherlands, Ireland and Bulgaria. The project's main purpose in Bulgaria was to test a new model for school inspection – innovative for Bulgarian inspection practice – to inspect a network of schools, conceptualized in the framework of the project as a form of 'polycentric inspection' (Simeonova and Parvanova 2017a, b).

Polycentric approaches to school evaluation take different forms in various social and regulation settings. The Bulgarian experience in the application of this model provided an opportunity to reconsider existing procedures and practices for school inspection and evaluation, and to realise more up to date school evaluations, relevant to the current conditions and needs of the local education context.

Within the project's framework a successful cooperation of three different types of institutions was accomplished – schools, the inspectorate and a university, collaborating to improve the quality of education and school management in Bulgaria. The research team from the Faculty of Education at Sofia University "St. Kliment Ohridski" coordinated the project activities in Bulgaria. The Sofia Regional Inspectorate of Education (SRIE) was invited to be a project partner. Finally, 10 schools from Sofia (4 primary and 6 comprehensive schools from different city districts and with various profiles regarding achievements, experiences and challenges they are facing) were also invited to join the project. All school principals were recognised professionals and highly motivated to test innovative practices. The invited schools established a voluntary network to cooperate on testing a 3-stepped polycentric evaluation model. The model was designed by the Sofia University research team integrating self-evaluation, peer-evaluation and inspection of schools collaborating in a non-formal network (Simeonova and Parvanova 2019a, b).

For the purposes of the case study presented, peer-evaluation/peer-review could be considered a form of internal or external school evaluation, undertaken by evaluation teams from other schools to review educational and/or managerial practices and achievements of a particular school. At network level it is an internal one, but from individual schools' perspective it is an external one. Although a peer-review could be a form of internal evaluation as it doesn't involve an authority in evaluation process (Nelson et al. 2015), the case-study presented here is also in line with the notion of external evaluation of the school. It has been preceded by a school self-evaluation and has been followed by a polycentric inspection. As peer-review was part of a three-step process, its design followed the design of an external evaluation without the high stakes accountability usually related to it. Such types of evaluation are often perceived more positively in comparison to other forms of external school evaluation (i.e. inspections), as evaluated parties see a lot of benefits for their professional development and are usually highly satisfied by its supportive format.

6.2 External and Internal Evaluation of Bulgarian Schools

An overview of the Bulgarian school system is given in Box 6.1 below for those not familiar with this national context. Focusing on school evaluation and accountability, the bodies responsible for external school evaluation, established under the new Preschool and school education act (2015) are:

- *National inspectorate of education (NIE)* – subordinated to the Council of ministers and independent from the Ministry of Education, this body conducts whole school inspections in every school every 5 years or more often depending of the

Box 6.1: Bulgarian Educational Context
Bulgarian school education system encompasses 2 years of obligatory preschool education (age 5–7) and 12 years of school education (age 7–19). Preschool education is provided by both kindergartens and schools, but most children attend preschool classes in kindergarten. Obligatory school education is up to 16 years of age.

The school education system consists of two levels – primary (1st–7th grade) and secondary (8th–2nd grade) education. Each level is divided in two stages. Primary education consists of two stages – elementary (1st–4th grade) and middle school (5th–7th grade). Secondary education includes two stages – 8th–10th grade and 11th–12th grade. 10th grade finishes with certificate for those students who don’t want to proceed to the second stage of secondary school. Graduation at the end of 12th grade provides students with a secondary school diploma which allows them access to higher education. There are national exams and standardized testing at the end of 4th, 7th and 12th grades.

The School education system is managed by the Ministry of Education and Science (MES), *28* Regional Departments of Education. The RDE is subordinate to the MES and is responsible for the 28 regions of Bulgaria, former Regional inspectorates of education and municipalities.

School governance is highly centralized. In general, school autonomy is rather low, but financial autonomy of the schools is high as well teachers’ autonomy to choose teaching methods and resources. School principals are appointed by RDE and teachers are appointed by school principals.

At **school level managing bodies** are:

- *the principal;*
- *deputy principal(s);*

(continued)

Box 6.1 (continued)

- *pedagogical council* (consisting of all teachers and other specialists in the school, i.e. school counsellor, etc.; makes decisions on all strategic and important school issues, votes school's development plan, year plan, etc.);
- *public/community council* (consists representatives of parents, teachers, local public figures, other partners of the school).

The majority of the schools have a students' self-management bodies – '*students' parliament/council*' consisting of representatives of students from each class usually from the secondary school and sometimes from the middle school. This body discusses, plans and makes decisions on students' issues and has representatives in the pedagogical council in case students' matters are being decided.

The majority of the schools have a *Board of Trustees*. Traditionally in most Bulgarian schools such a body exists. It is a legal voluntary body, consisting parents and other partners of the school, established to support particular school but without managing authority.

In all Bulgarian schools there is a decades long tradition of class-based *parental councils* (usually consisting of 3 parents) to support class tutor and class activities.

inspection results. It does not execute support functions regarding inspected schools (Council of Ministers 2018; Ministry of Education and Science 2016a).

- *Regional Departments of Education* (RDE) (former Regional Inspectorates of Education), responsible for controlling and supporting schools within their regions and for organizing the national exams (Ministry of Education and Science 2017).

Prior and during the testing of the polycentric evaluation model in Sofia city *28 Regional Inspectorates of Education (RIEs)*, subordinated to the Ministry of education and science, were executing control over schools within their regions through different types of 'checks' (Ministry of Education and Science 2003). These RIEs later became Regional Departments of Education (RDE). Functions of RIE involved: whole single school inspection, thematic inspection (on particular areas of education provided by a number of schools in the region), ongoing inspection (of a principal or a teacher), and incidental/risk based inspections (in case of violation of educational regulations any stakeholder could notify the RDE which is obligated to investigate and make recommendations). These bodies were also entitled to provide support to the schools in their efforts to improve local educational provision. The polycentric inspection of the Sofia network of schools was at the time of its implementation and still is, highly innovative practice for the Bulgarian inspection context.

Internal evaluation is implemented by schools through assessment of their students' achievements, appraisal of teachers' performance and occasionally in some schools through a more thorough process of self-evaluation. The Preschool and

School Education Act states that school self-evaluation is obligatory for all schools and more details are specified in the document *Regulations for Managing Quality in (Educational) Institutions* (Ministry of Education and Science 2016b). The legislation was applied in December 2016 but eventually suspended in December 2017 due to resistance by schools and a lack of methodological support provided to schools for the implementation of self-evaluation. According to this regulation schools were obligated to develop and implement self-evaluation criteria and instruments, although no further guidelines were provided. Most principals and teachers didn't have self-evaluation competencies or prior experience with it and didn't feel confident to implement self-evaluation of their schools, so they made a strong case against the regulation. Self-evaluation of the schools within the Sofia network was implemented voluntarily in January 2016 as part of the Erasmus+ project prior to application of these legislative changes. The majority of the schools within the Sofia network were also lacking competencies in self-evaluation and didn't have any in peer-evaluation. Therefore, Sofia University provided this training to the principals. They felt the polycentric approach to school evaluation (implemented in the Sofia network of schools as self-evaluation, peer-evaluation and network inspection) would be beneficial to them in terms of gathering relevant internal and external evaluation experience and to develop these skills.

Regarding school networks in Bulgaria, there are a few examples of functioning school associations where schools cooperate and share good practices, for example some associations of vocational schools, private schools or the Association of Cambridge schools (mostly of municipal schools) which provide English language education following the Cambridge language teaching methodology. The establishment and functioning of non-formal, voluntary networks of schools, for the purposes of collaborating to exchange good practices, addressing common challenges, testing new models and practices, and creating joint products and know-how, are extremely rare in the Bulgarian educational context.

School peer-evaluation is not regulated by educational legislation in Bulgaria, nor is it practiced, besides within-school peer review between teachers and students. Hence our Sofia network case study was particularly innovative.

6.3 The Peer Review Model

During the year following the establishment of the Sofia network of schools (09.2014–08.2015) the schools met regularly to share good practices on topics and issues of common interest regarding the education they provide. Parental involvement was one of the common issues recognized by all partnering schools as needing improvement. Other activities of the network included designing tests for evaluating students' achievements at different educational levels, training seminars on school self-evaluation and school strategy development. Although some of the schools were collaborating occasionally prior to joining the network, the schools in the network gradually got familiar to each school's context and achievement through the practice of each network meeting to be hosted by a different school. SRIE representatives and researchers from Sofia University participated in each network meeting, providing expertise and methodological support. Sofia University coordinated these meetings and network activities. This phase was crucial for building up partnership, trust, and a sense of belonging to the network.

As result of these activities, at the beginning of the second phase of collaboration of the schools within the network (09.2015–08.2016) a framework for self-evaluation and peer-evaluation of parental involvement was devised. This framework was designed by the principals from the network with the methodological support of Sofia University research team. The School Review Guidebook of the National Association of Head Teachers in England (NAHT 2014) was used as a methodological basis. The appendices to the framework (instruments/forms for gathering evaluation data) were designed by the Sofia University research team.

The framework consists of:

- A definition of quality parental involvement in school
- Four standards for quality parental involvement in school
- Indicators for parental involvement for each standard.
- Sources of information.
- Evaluation methods
- Evaluation instruments
- A Scale for valuing and judging school performance on parental involvement
- The Self-evaluation/peer-evaluation period
- Data collection instruments
- A school self-evaluation report format
- Peer-evaluation report format.

For the purposes of self-evaluation and peer review within the Sofia network the definition of quality parental involvement was developed as follows: *“Quality parental involvement in schools includes various forms of effective communications, active participation of parents in school life and parents’ participation in decision-making regarding school development in order for the school to achieve its educational goals.”*

The framework for self- and peer-evaluation included four standards for quality parental involvement in school:

- school-parents communication supports educational process and students’ progress;
- parents participate in school life;
- school-parents interaction contributes to students’ progress and to the sustainability of students’ success;
- parents participate in the process of setting priorities for school development and improvement.

Network partners elaborated indicators for each of these four standards, which served as a measurement benchmark for the fulfilment of the standard. For example, some indicators for standard one included: “the school includes parents in preparation and update of the school development strategy; on a yearly basis a parents’ survey is implemented for the purposes of preparation of school's year plan; the board of trustees publicises its initiatives and reports to parents and teachers on different forms of support it has provided; the board of trustees' chairman participates in the work of the Pedagogical Council", etc.

Various sources of information were agreed by all partners for each indicator. These included surveys of stakeholders' opinion – teachers, students, parents, management team, evaluating minutes of parental and other relevant meetings, questionnaires, and the school plan for parental involvement, other relevant to parental involvement documents (official letters and other forms of correspondence and means of communication).

Different evaluation methods were applied during peer-evaluation, including meetings and discussions with teachers, students, parents, management teams, and analysis of official school documents.

In order to implement self- and peer-evaluation in a most objective way a scale for valuing and judging school performance on parental involvement was developed as element of the framework. It was especially important for schools to have a clear view on how to judge and evaluate other schools in the network. The scale comprised of three dimensions (see Box 6.2) and included rating scales for each indicator, along with overall scale to judging each standard regarding the level of achievement of each indicator.

Box 6.2: Scale for Evaluation of School Performance in Parental Involvement – 3 Dimensions:

1. Evaluation whether indicators for all quality standards have been achieved and to what degree (not achieved, achieved to a medium degree, achieved to a high degree).
2. There are minimum requirements for achievement of each of the four standards (1 or 2 particular indicators should be achieved at least at a medium degree).

Rating-scale:

- The standard is not accomplished if the minimum indicators are not achieved.
- Satisfactory – only minimum indicators are achieved.
- High – more than the minimum indicators are achieved.
- Excellent – all indicators are achieved.

3. Evaluation of quality of parental involvement in school – overall statement. Rating-scale:

- Low quality – in case standard 2 and 3 are not accomplished.
- Satisfactory – in case only standard 2 and 3 standards are accomplished.
- High – in case more than the minimum standards are accomplished at least at satisfactory level.
- Excellent – all standards are accomplished.

The self-evaluation/peer evaluation period was the previous school year till the beginning of self-/peer evaluation (09.2014–12.2015).

As peer-evaluation was preceded by school self-evaluation, the evaluation framework included four data collecting instruments for gathering self-evaluation data, summarised in a self-evaluation report and provided to the peer-evaluation team prior to the peer review visit in the host school. These tools are:

- Questionnaire for self-evaluation of school management team.
- Questionnaire for self-evaluation of teachers.
- Questionnaire for parents.
- Questionnaire for students.

A peer-evaluation report form was devised and proposed by the Sofia University research team and subsequently discussed and agreed by the partners in the network. The purpose of this short report (averaging 8 pages) was to file and summarise all the peer review data gathered by the peer evaluation team and this was presented in a standardised format to the principal of the host school (see Box 6.3).

Box 6.3: Peer-Evaluation Report Structure

Introduction (evaluated school; peer-evaluation period; peer-evaluation team; participants from the school).

Findings (for each standard).

Summary of the findings (strengths; areas that need improvement).

Level of concurrence between the opinions of school management team, teachers, parents and students.

Conclusion (statement) about the quality of parental involvement in the school. This includes:

- Degree to which the standards are achieved.
- Quality of parental involvement in the school.
- Adequacy of the planned measures and activities for improvement (as described in the school self-evaluation report).

Recommendations

- What to be sustained as good practice of parental involvement.
- What to be improved and developed in the area of parental involvement.
- What support from the network school can rely on for implementation of the planned improvements

Date and signatures of peer-evaluators

School principal comments

Appendix. Summary of the information gathered through the peer review process – for each standard and indicator.

The last element of the peer review report was an appendix, providing a summary of the information gathered through the peer evaluation process for each standard and indicator provided in the 3 sections. The first section summarises the school self-evaluation findings, stated in the self-evaluation report – summarized information for the inputs, processes and outputs as main sources of the self-evaluation reporting: what has been planned regarding parental involvement as stated in the school year plan and school development strategy (input); activities implemented in relation to parental involvement during the evaluated period (processes); stakeholders' opinion (parents, students, teachers, school management team) for the implemented activities and their satisfaction by school–parents partnership (outputs). The second section summarises the peer review findings, based on 3 sources of information: school self-evaluation report and a critical analysis of the information it presents; direct impressions of the evaluation team during the school visit; and the information gathered through discussions with school management team and representatives of teachers, students and parents (class-based parental groups, board of trustees) in order to verify and supplement the information provided in the self-evaluation report. The third section was 'comments' where the peer evaluation team could present a brief rationale that clarifies or underlines their judgments as well as some consideration taken into account like external context, specific circumstances, socio-economic profile of students, or the school's profile.

6.4 Peer Evaluation Procedure

Peer-evaluation was implemented within a 2-week timeframe (29.02–12.03.2016). Each school was visited for a day by an evaluation team, consisting of representatives of two other schools of the same type from the network (primary or comprehensive) – a principal, deputy principal and/or 1–2 teachers. The evaluation team examined the self-evaluation report of the evaluated school prior to the school visit as preliminary data and as a basis for gathering further evaluation data during the visits. The evaluation team held meetings and discussions with school leadership teams, teachers, students (representatives of student councils/parliaments) and parents (representatives of class-based parental councils and school board of trusties). They also scrutinised other school records and documentation relevant to parental involvement. After their meetings with all stakeholder groups the peer-evaluation team provided preliminary feedback to the principal of the host school on site.

The report for each school was produced by the evaluation team consisting of evaluators from two other schools. The evaluators divided among them the four parental involvement standards to come up with a statement for the level of achievement of each of the standards, then discussing the findings together and agreeing on the final version of the peer review report to be provided to the principal of the host school (and subsequently to the teachers by the principal and to the parents through various communication channels). Once receiving the report, the principal of the host school could add comments to which recommendations he/she accepts and which he/she disagrees, although in practice none were forthcoming.

For each indicator and standard, the peer-evaluation team assessed to what degree it was achieved and stated arguments to support the judgements, referring to relevant sources of information used for the valuing. Based on this procedure an overall statement for the quality of parental involvement in the host school was produced by the evaluation team by applying the rating scale presented in Box 6.2 above. While preparing the overall statement for the quality of the parental involvement in the host school, the peer-evaluators also considered the school's external context factors which might influence school performance, such as students' characteristics, parents' characteristics, the school's budget, its traditions and community and municipal support.

Overall peer-evaluation findings, benefits and positive effects were discussed and recognized by all partners (Sofia school network, Sofia Inspectorate of education, Sofia University research team) during a closure meeting (28.03.2016) and the follow up polycentric inspection of the network was planned for the next month, i.e. in April 2016. The inspectorate developed a framework for inspection of the network on parental involvement based on the self- and peer-evaluation framework designed by the schools and took into account self-evaluation and peer review reports while making judgements and in preparation of inspection reports. This approach followed a constructivist and flexible method to the network's inspection by adapting the network's own evaluation criteria and consequently producing inspection statements and recommendations more relevant to the network's agenda and more useful for the schools.

6.5 Case Study Methodology

The overall case study approach as part of the Erasmus+ project mentioned above was a longitudinal case study, implemented in 4 steps: after self-evaluation and before peer-evaluation, after peer-evaluation and before inspection, after the inspection of the network, 8 months after the polycentric inspection. For the purposes of this chapter and analysis of the findings discussed below, only the research data gathered right after the implementation of the peer review within the Sofia school network will be presented.

The ***research methods*** include action research, interviews and focus groups with network principals, questionnaires for evaluation teams' members and teachers from the network, observation of the peer review process/procedure, documents analysis and desk research (see Table 6.1). The school peer evaluation visits were shadowed by the Sofia University research team and observation notes are considered in the following analysis of findings. **Two main variables** were considered for

Table 6.1 Peer review case study phases, sample and research methods

Case study phase	During the peer review	After the peer review	
Timeframe	*29.02–12.03.2016*	*End of March 2016*	
Sample	***Methods***		*Research instruments designed by:*
10 principals	Shadowing	Interviews Focus group (28.03.2016)	E+ international research team and adapted for Bulgarian context by Sofia University Research Team (SURT)
Peer-evaluation teams (27 respondents)		Questionnaire	E+ international research team and adapted for Bulgarian context by SURT
158 teachers		Online survey	SURT
	Documents analysis School development strategy, school year plan, other mandatory documents required by educational legislation	**Desk research** Peer evaluation reports Observation minutes Focus group minutes	SURT and agreed within the network

the purposes of the national case study within the framework of the project, relevant to peer-evaluation data gathering and analysis:

Peer-evaluation practices in the network – peer-evaluation framework and procedure; peer-evaluation process: visits and follow-up.

Effects and benefits of peer-evaluation within the network – this included potential dysfunctional effects, i.e. power struggles, conflicts, changes of communication pattern and trust among network members, decreasing freedom of expressing different points of view, decreasing motivation of school staff to participate in the network as well as the positives effects for the network.

The network's characteristics and functioning, and the external context of the network were also studied as variables, but do not form part of this chapter.

The sample included a total of 195 respondents, 10 of them principals, 27 members of evaluation teams (10 deputy principals, 5 school counsellors, 12 teachers) and 158 teachers from the schools. Of all teacher-respondents in the survey, 90% were women, 44% teaching at primary level, 32% at middle school level and 24% at secondary level of education. 69% of teachers were class tutors who would have frequent contacts and interactions with parents.

6.6 Findings and Discussion

6.6.1 Peer-Evaluation Practices in the Network

6.6.1.1 Peer-Evaluation Framework and Procedure

Most members of the evaluation teams were positive about the peer-evaluation framework; it was described by one team member as "clear, effective, exact, well-structured, easy to use, and providing a realistic notion of parental involvement in school". Half of the principals pointed out as a positive fact that the framework had been developed by the network and believed that this made its implementation easier and supported evaluation teams in taking schools' specific circumstances into account. Members of evaluation teams described the peer-evaluation framework as "adequate" and "effective", while teachers believe it supported interaction and exchange of experience between schools to a great extent.

Some difficulties in framework application and some opportunities for improvement were pointed out by the members of evaluation teams such as: some indicators need to be more clearly differentiated; the rating-scale should be more precise as this will ensure considering schools' specifics.

Most members of evaluation teams (85%) described the peer-evaluation procedure positively as *suitable, well-structured and organized, clear, and providing opportunities for achieving peer-evaluation goals*. Evaluation team members said that evaluation of small schools by big ones and vice versa was a positive experience, providing opportunities to gather information about different types of schools which are in different positions within the Sofia school system.

Some members of the evaluation teams pointed out some barriers and obstacles due to the novelty of peer-evaluation procedures in Bulgaria. These included:

- lack of time or additional resources allocated for school visits and the peer review process;
- lack of standardised forms for gathering and filing evaluation data to be used during the visits by the peer-evaluation team;
- the number of students and parents to participate in the meetings with the evaluation team was not decided prior to the peer evaluation procedure;
- difficulties organising teachers and parents to participate in the meetings with the evaluation teams
- some concerns about the preparation of the peer-evaluation report;
- the organizational load for evaluation teams;

6.6.1.2 Preparation for the Peer Review Visit

Principals who participated in peer-evaluation teams elaborated the various forms that preparation for peer review visits took, this included reviewing school self-evaluation reports and discussing the organization and distribution of responsibilities within the school team. Parents and students had been invited to the meetings beforehand and also teachers to participate in the conference with the evaluation team. Documents and artefacts that could validate statements made in self-evaluation report were also available for the purposes of the visits. Most members of evaluation teams shared that they had been engaged with the organization of their school peer-evaluation, and with communication with parents, and students for the forthcoming meetings. The majority of the preparation for the peer-evaluation consisted of planning these meetings.

99.4% of all teachers stated they were aware that a peer-evaluation of parental involvement had taken place in their school. 50% of them participated in the self-evaluation of their own school or participated in the meetings with the peer-evaluation team during the peer review of their schools.

Principals shared some ideas for improvement of preparation of their school for peer-evaluation. These included: making evaluation criteria more precise; clarifying procedures for inviting different stakeholder groups to the meetings with the evaluation teams; building in time for the school to present itself fully; and preparing example sets of documents to support evaluators' judgments. Overall, peer-evaluation was perceived mostly positively, with the caveat that it took time to become used to the organizational and procedural aspects that are required to make them function perfectly.

6.6.1.3 Peer-Evaluation Process – Visits and Follow-Up

The majority of evaluation team members (80%) believed that the school's self-evaluation had been taken into account to a high extent during the peer-evaluation process and agreed that the process was "complete, clear, organized and exact, and provided the opportunity for a quality exchange of information". Principals appreciated that peer-evaluators were other principals and teachers, and representatives of professional school communities. They felt that these stakeholders had good intentions and were willing and able to identify the strengths of the evaluated school.

During the study, both principals and other evaluation teams' members stated that peer-evaluation feedback would be presented (or had already been presented) in summary to the Pedagogical Councils of their schools (see Box 6.1 for the functions of that body). Schools were planning to present the results also to the board of trustees, and 4 schools to the students' parliaments/councils. In order to publicise the peer-evaluation findings respondents also considered opportunities to upload the peer-evaluation reports (full or summary) to the school internet site, to make it widely accessible and subsequently some of them did so.

According to 61% of the teachers who participated in the survey said that peer-evaluation results had been already presented to the pedagogical staff, and they had had the opportunity to discuss them, and to outline possible measures for improvements. All schools were planning to present these results to the school community in one way or another. School principals and other members of the evaluation teams were determined to present peer-evaluation results to the parents and subsequently some of the schools did so using the school's website and Board of trustees to disseminate report findings. Six out of ten schools were also planning to prepare a summary of the peer review findings and present it during a parents' meeting.

All peer-evaluation visits were shadowed by the research team from Sofia University, providing the research team with data to summarise some positives and difficulties of their implementation:

Positives:
- All schools had been prepared for the evaluation. They presented self-evaluation reports to evaluation teams and organised meetings with parents, teachers, and students. Evaluation teams used these to prepare additional questions to take to the review visit.
- The atmosphere at the schools during the peer evaluation visit was positive and calm in all cases and the evaluators' approach was considered good, stimulating beneficial discussions.
- School staff and evaluation teams shared experiences not only about the evaluated area but about other areas of schooling as well.

Difficulties:
- On some occasions the lack of experience in peer-evaluation created difficulties for evaluation teams, in particular how to ask adequate and focused questions in order to receive deep and complete information on the evaluated topic.
- The evaluators' status of 'colleagues' to the principals and teachers in the evaluated schools sometimes made it difficult to be objective evaluators and/or to provide the appropriate challenge.
- In certain situations, the exchange of experience about good practices prevailed over data gathering in the evaluated area – parental involvement.
- Sometimes smaller schools felt uncomfortable evaluating bigger schools and schools in city centres with good reputations (highly performing schools).
- Evaluated schools did not provide additional documents to support and verify their self-evaluations therefore evaluators relied on the data gathered through discussions with parents, teachers and students.
- At the closure of the school visits some evaluation teams did not provide preliminary feedback to the principals because they felt uncomfortable and inexperienced in doing this.

By working on the project, principals gathered a broad range of knowledge by studying evaluation frameworks, participating in trainings and seminars and sharing good practices. However, some participants in peer-evaluation teams felt they did not have enough skills and confidence to evaluate their colleagues' work.

6.7 Effects and Benefits of Peer-Evaluation Within the Network

6.7.1 Potential Dysfunctional Effects

Most evaluation teams' members shared the opinion that freedom to express different points of view in the network had not changed since the establishment of the network. This freedom was already very strong and was not negatively influenced by the peer-evaluation. Their positive experience of peer-evaluation helped free them of concerns about how they as professionals and the work of their schools might be judged by other members of the network. It also strengthened their feelings of acceptance and understanding with network partners. This is especially evident for deputy principals and teachers, participating in evaluation teams. Most principals and evaluation teams felt that there had been equal participation by all partners in the network.

Teachers were asked whether they believed if their colleagues' motivation for participation in both the network and the project had declined, compared to the one in the beginning of the project. 47.5% of them answered "No", but almost 35% stated that they were undecided on this issue. Principals and other members of evaluation teams were quite confident that no partners would leave the network; there was a common sense of satisfaction with the work done and from the planned further network activities.

6.7.2 Positives of the Peer-Evaluation

Peer-evaluation of schools in the network was expected to bring some beneficial effects for both schools and the network. Being evaluated by peers should provide schools with useful feedback about the way they performed in terms of parental involvement. On the other hand, communication in the network should benefit as more school staff would be able to visit and communicate with their peers in other schools in the network. Peer-evaluation visits would also provide a chance to see and experience other schools' practices and this would support the exchange of ideas and working methods between partners in the network. All of the above were borne out through the data collected in this case study.

All evaluation team members shared in the questionnaire the same opinion that peer-evaluation feedback was extremely useful. Some of them commented that it was incredibly enriching to share experience with their peers, while having an external point of view for their work. This provided an opportunity for self-reflection and comparison with other schools. Receiving constructive feedback on the weaknesses of parental involvement activities was especially beneficial for planning relevant improvements. In the interviews, principals shared that changes and improvements had already been planned in their schools as a result of peer-evaluation and the feedback it provided. All teachers and school counsellors also thought that such changes were planned as result of peer-evaluation. Most frequently pointed out planned changes, were: developing and implementing schools' trainings for parents in their schools; special meetings with parents to plan different school activities; introducing protocols for parental conferences (taking minutes, etc.); and more effective promotion of school activities for parents to participate in. After the peer-evaluation most of the schools made improvements regarding parental involvement by adopting good practices from other schools in the network. In one school, practices observed in a reviewed partner school were introduced, including improving communication with the school Board of Trustees, while other schools implemented new forms of parental meetings, employing some best practices they had seen in the network.

All members of the evaluation teams, with some minor exceptions, shared that ideas for further cooperation between schools in the network had already emerged. These included: meetings and joint activities of boards of trustees and students' parliaments, participation of the network in different projects, developing joint initiatives and sharing their network experience with other schools.

Principals and other members of evaluation teams found shared experience and comparison with other schools as the main positive of their schools' participation in the peer-evaluation, along with feedback by colleagues from different schools. Teachers who participated in the survey after peer-evaluation (total of 158 teachers from all 10 schools) shared similar views to the evaluation teams. Results show that approximately 2/3 of all surveyed teachers believed that the peer-evaluation was useful (see Fig. 6.1).

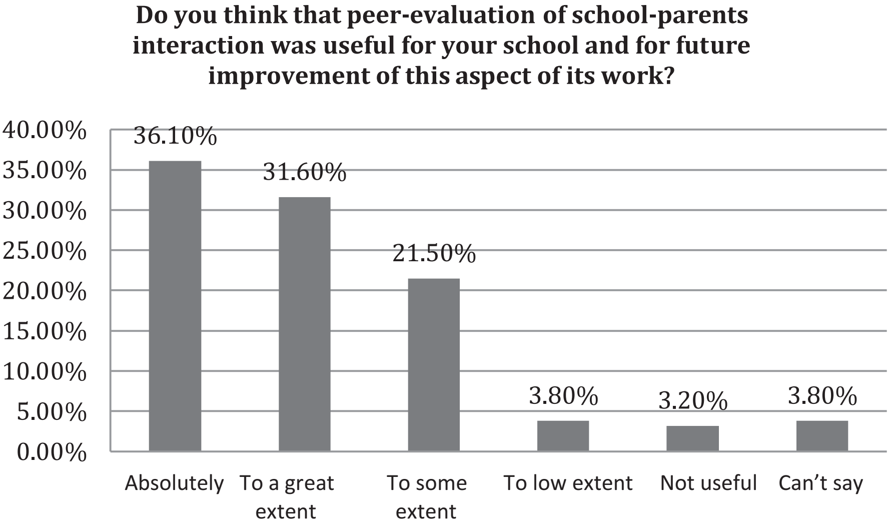


Fig. 6.1 Teachers' opinion about peer-evaluation benefits for their school

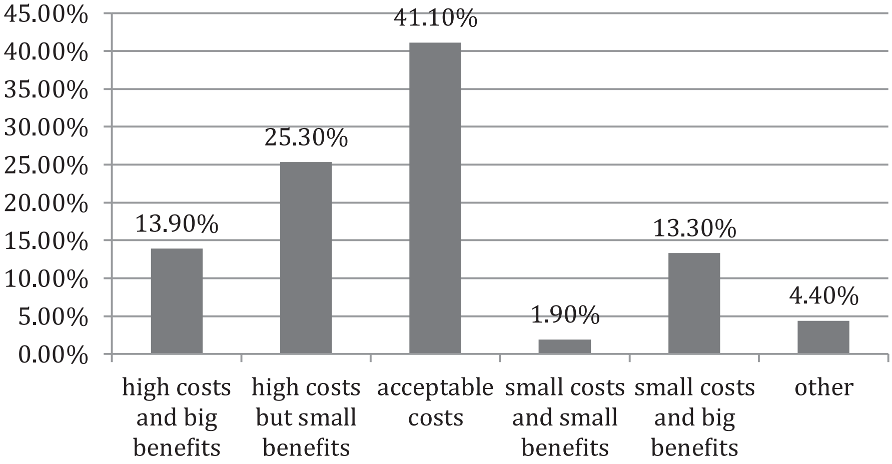


Fig. 6.2 Teachers' opinion about peer-evaluation costs-benefits proportion

Meanwhile, teachers' opinions about the costs-benefits ratio is especially interesting (see Fig. 6.2). Data show that 2/3 believed that peer-evaluation benefits were great, and for 41% of teachers the costs were acceptable.

From the data gathered from evaluation teams and teachers it is evident that participation of schools in the peer-evaluation was mostly perceived as positive, with further positive effects on schools, despite being somewhat difficult and time-consuming. The **main benefits** pointed out by the principals are related to the improvement of their knowledge and skills related to managing parental involvement. In addition, participation in the peer-evaluation improved their competencies to reflect on and to assess their own work, the work of their employees, the whole school work and in a certain area of schooling. The main professional benefits were exchanging experiences, comparison with colleagues and reflection on the way they performed their professional duties. Especially useful for deputy principals, school counsellors, and teachers was the opportunity to establish professional contacts with other colleagues in the same position in other schools and to discuss common issues within the context of their collaboration with parents.

When asked whether they planned any changes in school work as a result of network participation, principals provided quite detailed answers, stating at least one or two ideas for such changes, most of them related to parental involvement. These included adding specific duties regarding parental involvement in teachers' job descriptions and using more electronic forms for communication with parents and for presenting the school's achievements to the public. In addition, strengthening the work with the Board of Trustees and more engaged project work were also considered. Other members of evaluation teams suggested developing and providing training sessions for parents along and more regular surveys of parental opinions, employing more electronic forms of communication.

6.8 Conclusions and Follow Up

Peer-evaluation in a network of schools was a completely innovative practice in the Bulgarian education context and was implemented in a legislative ‘vacuum’. All principals and evaluation team members expressed their high satisfaction with the knowledge and experience acquired, despite the difficulties they faced with their new role as peer-evaluators. The fact that the Sofia Regional Inspectorate of Education considered the peer-evaluation data in making their polycentric inspection judgements is especially valuable.

The future of peer review of schools in Sofia seems optimistic having in mind the positive feedback provided by all participants, the benefits outlined by the schools involved and continuous collaboration of the network partners. One sign of the successful cooperation of the partners in Sofia network of schools was their motivation to keep working together. In January 2017 they established an NGO called the Network for Innovations in Education. These schools continue to cooperate on other issues and topics of mutual interest after the end of the project, to share experience and know-how with other principals, educators and other interested parties. For instance, they have chosen the topic for next peer review to be implemented within the network – approaches and practices for students’ assessment.

In order to be disseminated further and evolve into more sustainable practice the peer-evaluation of schools in Bulgaria would need a more supportive infrastructure involving additional methodological support by the universities, more dissemination and training events, changes in the relevant regulations to stimulate schools to test new practices, and governing bodies to support schools’ innovative developments so they could benefit all stakeholders.

References

Council of Ministers. (2018). *Regulations for the structure and functions of the National inspectorate of education* (Pravilnik za ustrojstvoto i funkciite na Nacionalnija inspectorat po obrazovanieto), State Gazette, No. 36/13.03.2018. https://dv.parliament.bg/DVWeb/showMaterialDV.jsp?idMat=124209

Ministry of Education and Science. (2003). *Regulations for the structure and activities of the regional inspectorates of education* (Pravilnik za ustrojstvoto i dejnostta na regionalnite inspectorati po obrazovanieto), State Gazette, No. 61/8.07.2003. https://dv.parliament.bg/DVWeb/showMaterialDV.jsp?idMat=109101

Ministry of Education and Science. (2016a). *Regulations for inspection of kindergartens and schools* (Naredba No. 15 za inspektiraneto na detskite gradini i uchilishtata), State Gazette, No. 100/16.12.2016. https://dv.parliament.bg/DVWeb/showMaterialDV.jsp?idMat=110175

Ministry of Education and Science. (2016b). *Regulations for managing quality in educational institutions* (Naredba No. 16 za upravlenieto na kachestvoto v instituciite), State Gazette, No. 100/16.12.2016, suspended with order of the minister of education, State Gazette, No. 100/15.12.2017. https://dv.parliament.bg/DVWeb/showMaterialDV.jsp?idMat=110176

Ministry of Education and Science. (2017). *Regulations for the structure and functions of the regional departments of education* (Pravilnik za ustrojstvoto i funkciite na regionalnite upravlenija na obrazovanieto), State Gazette, No. 13/7.02.2017. https://dv.parliament.bg/DVWeb/showMaterialDV.jsp?idMat=111568

NAHT. (2014). *School Review Guidebook*. National association of head teachers in England.

Nelson, N., Ehren, M., & Godfrey, D. (2015). *Literature review on internal evaluation*. London: University College London, Institute of Education.

Preschool and School Education Act .(2015). (Zakon za preduchilishtnoto i uchilishtno obrazovanie), State Gazette, No. 79/13.10.2015, applied on 01.08.2016. http://dv.parliament.bg/DVWeb/showMaterialDV.jsp;jsessionid=0CE25EBBCEBF8D8E0031EFAD78701AE4?idMat=97877

Simeonova, R., & Parvanova, Y. (2017a). Innovative model for polycentric inspection of network of schools in Bulgaria (Inovativen model za policentrichno inspektirane na mreja ot uchilishta v Bulgaria). In *Annual book of Sofia University 'St. Kliment Ohridski', Faculty of Education, Education* (Vol. 110, pp. 90–119). Sofia: St. Kliment Ohridski University Press. https://www.uni-sofia.bg/index.php/bul/universitet_t/fakulteti/fakultet_po_pedagogika/oficialni_izdaniya.

Simeonova, R., & Parvanova, Y. (2017b). *Polycentric inspection of network of schools: A case study from Bulgaria.* http://www.schoolinspections.eu/

Simeonova, R., & Parvanova, Y. (2019a). *Evaluating schools collaborating in a network: Experience from innovative practices in Bulgaria* (Ocenjavane na uchilishta raboteshti v mrezha: opit ot inovativni praktiki v Bulgaria). Sofia: St. Kliment Ohridski University Press.

Simeonova, R., & Parvanova, Y. (2019b). Peer evaluation of schools collaborating in a network - case study results and innovative practice (Vzaimno ocenjavane na uchilishta, raboteshti v mrezha –rezultati ot prouchvane i inovativna praktika). In *130 years of university pedagogy*. Sofia: Prosveta Publishers.

Dr. Rossitsa Simeonova is Associate Professor at Sofia University "St. Kliment Ohridski", Faculty of Education. Her research interests and teaching are related to educational management, educational evaluation and social work management. Principal investigator for Bulgarian portion of the Erasmus+ projects 'Polycentric inspection of network of schools', 'Supervising schools in the twenty-first century: digital tools and improvement plans' and 'New approaches in inspection: a polycentric model'.

Dr. Yonka Parvanova is Associate Professor at Sofia University "St. Kliment Ohridski", Faculty of Education. Her research interests are related to new paradigms and approaches to management of school systems and organizations, interdisciplinary aspects of financing and education system and achieving effectiveness and efficiency of the education system. Investigator for the Bulgarian portion of the Erasmus+ projects 'Polycentric inspection of network of schools' and 'Supervising schools in the twenty first century: digital tools and improvement plans'.

Chapter 7
Peer Review in Czech Education: A Recognized but Somewhat Neglected Tool for School Development

Stanislav Michek, Martin Chvál, and Milan Pol

Abstract This chapter presents the practice of peer review as related to self-evaluation processes in Czech schools. It begins with a description of changes that have occurred in self-evaluation of Czech schools during the last 15 years. This is followed by discussion of efforts to support peer review in the last hundred years and a description of how peer review was supported in the Road to Quality, a nationwide project (2009–2012) that approached peer review as evaluation by teams from other schools. Experience from the project is presented along with research methods applied, such as questionnaire survey, semi-structured interview, document analysis, participated observation, inquiry and case study. The findings give evidence on (a) expectations of schools willing to carry out peer review; (b) evaluation activities and work with school data; (c) differences between peer review and action of Czech School Inspectorate; (d) willingness of schools to continue using peer review for further development. In conclusion, peer review in the Czech Republic is perceived as a potentially promising activity; in terms of practice, however, it is shown that it remains somewhat peripheral in the mainstream of Czech schools.

7.1 Introduction

This article describes efforts to support the practice of peer review in Czech schools, which so far has tended to be applied experimentally and as an innovation, albeit over several decades. We will comment on the relationship between peer review and

---

S. Michek (✉)
Faculty of Education, University of Hradec Králové, Hradec Králové, Czechia
e-mail: stanislav.michek@uhk.cz

M. Chvál
Faculty of Education, Charles University, Prague, Czechia

M. Pol
Faculty of Arts, Masaryk University, Brno, Czechia

© Springer Nature Switzerland AG 2020

D. Godfrey (ed.), *School Peer Review for Educational Improvement and Accountability*, Accountability and Educational Improvement,
https://doi.org/10.1007/978-3-030-48130-8_7

self-evaluation, specifically focusing on the *Road to Quality,*[1] an important national project that supported evaluation actions in Czech schools and, among other initiatives, validated a peer review process for schools. We will describe the experience gained from efforts to develop and support peer review in schools and show that, despite all these efforts, peer review remains a potentially strong but mostly complementary practice in the mainstream of schools in the Czech Republic.

7.2 The Changes in School Self-Evaluation in Czech Education

The intense discussions about ways to ensure quality in a decentralized schooling system under new social conditions in the 1990s resulted in the origin of a strategic document, the *National Education Development Programme* (2001), also called the *White Book*. This document stipulated the need to partly transfer quality assurance to schools, including its control by school leaders. In the spirit of this logic, the duty to carry out their own evaluation was imposed on Czech schools by the School Act (2004/561), obliging them to self-evaluate, work out self-evaluation reports, file the reports as compulsory documentation and submit them to the Czech School Inspectorate (CSI). This measure was compulsory for nursery schools, primary schools and secondary schools.

It soon turned out that the absence of a working support system for self-evaluation in schools was a serious problem, which, among other consequences, resulted in low acceptance of the duty of self-evaluation. Thus, the point of school self-evaluation as a desirable tool of education policy was questioned by some of those involved. The situation evolved so far that after the Czech parliamentary elections of 2009 the obligation to produce a self-evaluation report appeared in the list of "administrative overload of schools"; in 2011, this obligation as well as the decree on school self-evaluation in its entirety were cancelled. Therefore, schools are still supposed to carry out self-evaluation, but it is in a rather loosened mode. Certain pressure on its quality can be made by the Czech School Inspectorate, but inspectors get to schools once in approximately 6 years only, in the form of a comprehensive inspection visit. In spite of this, the quality of self-evaluation is one of assessment criteria for the Inspectorate. Also, pressure on school leaders is made by the school establishing entities[2] in case they want to be regularly assured of the school's good performance.

In fact, the attitude of schools to self-evaluation varied greatly. On the one hand, there were committed schools for which the duty of self-evaluation did not make much of a difference, as they had been carrying it out on their own initiative for some time already. For these schools, it meant no more than legitimization of activities already existing. Headteachers and teachers at these schools usually became involved in associations of "active" schools (e.g. Friends of Engaged Teaching[3]). A 2009 questionnaire survey of headteachers (Chvál et al. 2012) showed that these schools had understood self-evaluation as a necessary component of their job and a part of their efforts to improve their work. However, these schools were not numerous, representing some 5% of the total only. At the opposite end of the scale were schools with an attitude of scepticism to self-evaluation, in most cases understanding it merely as another useless item of administration, and therefore a meaningless burden. Their leaders had a strong tendency to create the documents of self-evaluation in a formal manner, simply in order to meet inspection requirements, without relating these processes to efforts to improve their work. These schools were quite numerous, at almost 65%.[4] Between these two groups were active schools whose leaders were conscientious about the duty to carry out self-evaluation and willing to consider its importance. The leaders of these schools were able to acknowledge the benefit of internal discussions on school quality and regarded the writing of a report as a logical result of a relatively meaningful process. These schools amounted to almost 30% of the total. The committed schools (the first of the groups referred to here) did not need a support system. For schools rejecting the changes, however, the absence of a support system was an excuse for only formal fulfilment of requirements. The schools trying honestly to carry out self-evaluation felt the absence of a support system strongly.

---

[1] http://www.nuov.cz/ae?lchan=1&lred=1

[2] School establishing entities (e.g. municipality, private founder) is the founder of schools. In case of public school consists from politicians and school officers (administrative). Politicians at local level decide and officers assure administrative support.

The national programme the *Road to Quality*, meant to fill the gap in support for self-evaluating schools, was started as late as in 2009, some years after the need for such support emerged, as legislation had been altered in 2004. Ironically and rather by coincidence, effective assistance to schools started arriving only when further reforms to legislation were planned (2010–2011).

In spite of this, the *Road to Quality* was a substantial project within which peer review, among other activities, was implemented as we describe it below. The project profited greatly from cooperation with active schools and school leaders and their willingness to get involved in certain activities.

So, starting from 2012, schools in the Czech Republic were required to self-evaluate and project the results of their self-evaluations in annual reports on their work, these being made available to the public. As to external assessment, schools are evaluated by the CSI. The Inspectorate makes a general inspection visit to each school approximately once every 6 years; it carries out thematically focused surveys with selected samples of schools in between (e.g. performing research based on questionnaires for headteachers and student testing in certain school years). School evaluation by the Inspectorate includes assessment of the quality of self-evaluation based on observation, interviews with participants and document analysis. If the CSI came across peer review, no matter how rarely this can occur, they would probably consider it positive in terms of the criterion that says "School leaders actively manage and regularly monitor and evaluate the work of the school and take efficient measures".

---

[3] An interest association founded in 1992, looking for new and unconventional approaches to everyday activities in education and schooling.

[4] Expert estimation based on positive response to information project activities offered in a questionnaire (courses, workshops, conferences) and the low return rate of the questionnaire.

7.3 The History of Efforts to Support Peer Review in Czech Education

It is rather a matter of recent years that peer review has been mentioned as a tool for school development. Nonetheless, peer review activities such as visits to schools by teachers from other schools, sitting in on classes taught by one's peers and the exchange of experience among teachers from different schools — in other words and in a broader meaning, networking of teachers or schools so that they may learn from each other — are not completely new in Czech education.

There is a variety of evidence of similar initiatives, systematically organized, dating back to the era of the First Republic (1918–1938). Perhaps the most consistent programme was carried out by protagonists of the so-called reformative and experimental schools, e.g. in Prague's quarters of Nusle, Michle and Hostivař and the towns of Zlín and Humpolec, which started operating on 1st September 1929. These schools represented a reform movement conducted by the government and were coordinated by a reform committee presided over by Václav Příhoda. Called *Active Schools*, they accentuated autodidacticism and pupil testing and promoted project teaching. They claimed an affinity with Dalton Schools,[5] schools of the Winnetka Plan[6] and other elements which promoted reform pedagogy movement of the time. The activities of these schools attracted various visitors; their annual reports contain details of visits paid by representatives of the Ministry for Schooling and Education, important figures of public affairs, professionals from home and abroad, students of teaching institutes and, most importantly for peer review, teachers from many other schools in what was then Czechoslovakia.

At a time when reformative and experimental schools were coming to life, Votava (1928) and Koncer (1929) discussed the usefulness of sitting in on classes, which was rendered obligatory by paragraph 134 of the Order of Schools and Teaching. They considered mutual visits most suitable because,

---

[5] The Dalton Plan is an organizational form and method of teaching that was introduced to the world by Helena Parkhurst in the first two decades of the twentieth century. It is based on principles affecting the organization of teaching. The basic Dalton principles such as liberty, cooperation and independence invite pupils to become active participants of teaching. See more at Dalton International, https://daltoninternational.org/

[6] The Winnetka Plan is a modification of the Dalton Plan aimed at strengthening of group activities.

> "true democracy would dominate and the edges of sitting in would blur. Colleagues would learn from one another and improve for the benefit of pupils; there would be enough material for pedagogical debates at teacher conferences" (Koncer 1929).

Reciprocal classroom observations, with special attention paid to visits by experienced teachers in classes of novice teacher colleagues and the recommendation to visit each other as a "source of noble competition, particularly in methodology", is described by Bezdíček[7] (1938) in his Pedagogical Encyclopedia. He remarks that at technical schools, "novices who have not yet taught at any school should sit in on 4 to 6 classes per week taught by experienced colleagues, also in subjects other than their own, technical and theoretical, so that they may acquaint themselves with the procedure of teaching in the whole school". In his Reference Dictionary of Pedagogics, Pech (1937) adds: "Besides this meaning, 'sitting in on classes' describes reciprocal visits made by teachers, observations by trainee teacher and training-school teachers, and visits to other schools".

After World War II, particularly in the 1950's, the Soviet model of so-called pedagogical readings became widespread in Czechoslovakia (cf. Auerswald et al. 1957, Auerswald 1965, pp. 30–32, Peutelschmiedová 2011). These readings worked with teachers' manuscripts presenting practical inspiration (e.g. use of apiculture in schools, plant cultivation in nursery schools) and offered solutions for various issues in education (e.g. poor results, testing, marking). Teachers recorded their experience in the form of these small written works and used them in contests organized by various pedagogical institutions (e.g. Regional Pedagogical Institutes). Some of the more successful participants passed through various rounds of the contest (district, regional and, finally, national). They were rewarded with certificates, while articles written by the national-round winners were published in specialist journals (such as Pedagogická čtení [Pedagogical readings], 1956–1957, 1957). Besides professional work, participants in these contests were motivated by the possibility of career advancement, potentially becoming executives for schools administration or educational science.

A good example of enduring work with peer review after 1989 was the Programme for Quality Evaluation in Vocational Schools (EVOS), an initiative in vocational education; it is highly compatible with today's peer review methodology (Michek 2008). It was started in 1995 by the Association of Vocational Schools,[8] which was supported by the Ministry of Education, Youth and Sports. According to the Association's website (EVOS 2011), 20 vocational schools with 26 study programmes were evaluated by the programme in 1995–2000. The programme made use of the findings of self-evaluation during a follow-up process of external evaluation. It was based on discussions between the school and the evaluation committee on results, strengths and weaknesses as reflected by the self-evaluation report. Ideally, the committee consisted of a specialist in economics and/or social practice, a representative of an institute of higher education specializing in disciplines compatible with the programme under evaluation, a professional able to assess school management and development from the viewpoint of the post-secondary education system as a whole, and a representative of another vocational school specializing in comparable subjects of study. Evaluation within EVOS was focused on study programmes lasting a minimum of 3 years, assessing mainly its quality with reference to the overall concept and structure of the school. Having studied the self-evaluation report, the committee prepared an external evaluation including a visit to the school. Then the committee produced a report on the external evaluation and submitted it for comment to the school under evaluation. After incorporation of the comments and authorization of the results by the Managing Board of EVOS, the final wording of the report and the school's statement were published.

---

[7] Headteacher of *Městské české dívčí reformní reálné gymnasium* [Girls' Reformative Grammar School] in Brno, now Czech Republic.

[8] Now under the name Association of Professional Tertiary Education, http://www.ssvs.cz/

Similar examples can be found for schools with general education programmes, mainly at the level of the basic school (pupils aged 6–15). Within INOSKOP – Innovative Schools in Prague, a project with 12 basic schools in 2006–2008 (see Chvál and Starý 2008), 12 peer review plans of pedagogic development in schools were carried out.

In conclusion, peer review has been an occasional feature of Czech education but certainly not a systematic or widely used practice in mainstream schools.

7.4 Peer Review Within the Project *Road to Quality* in the Czech Republic

As mentioned above, one of the objectives of the project *Road to Quality* (2009–2012) was to support peer review activities in schools. In a certain sense, it was the most challenging of all the activities contained in the programme, or at least the most ambitious. The starting point for peer review implementation was the methodology for European Peer Review (Gutknecht-Gmeiner 2007), which was adapted in order to meet Czech conditions in 2009.

The objective of European Peer Review is to support institutions that are being evaluated, i.e. providers of vocational education wishing to enhance their quality. The procedure corresponds with the Common Quality Assurance Framework (Fundamentals … 2007) as created by the Expert Work Group for Quality in Vocational Education and approved by the Council of Europe in 2004. The main aim of this type of peer review, which has a formative function, is quality assurance in education. A major principle of European Peer Review was application of equal opportunities for men and women.

Peer review is performed by external groups of people who possess professional experience and occupy positions similar to those of their peers under evaluation. During evaluation, peers visit a school that — as a prerequisite — has undergone the process of self-evaluation at least once. Various stakeholders – schools, students, institutions, administration bodies, partners in vocational education and training, parents and prospective employers – participate in the procedure as members of the evaluating group. These stakeholders have not been trained in any way; they participated in peer review as respondents in questionnaire surveys, members of focus groups commenting on the quality of education provided, and so on. The peer review process is always identical, be it for two schools or a whole network.

Before the evaluation starts, the school decides on a particular domain of quality to be assessed. Peer review may be focused on as many as 14 areas of school operation, of which the first four are considered crucial for educational quality. These areas are: curriculum; learning and teaching; assessment; learning results/outcomes. The others are: social environment, access and diversity; management and administration; institutional ethos and strategic planning; infrastructure and financial resources; staff allocation, recruitment and development; working conditions of the staff; external relations and internationalization; social participation and interactions; gender mainstreaming; quality management/evaluation.

The term 'peer team' refers to the group of experts (facilitators) who perform the peer review. In our context, these teams consisted of Czechs independent of the school under evaluation, some with experience of a compatible type/level of school, others representing the labour market or other stakeholders. An important position in the team was held by an expert on evaluation, who was involved in the design of the process, feedback and, in some cases, data interpretation.

The adaptation of the tool to a Czech setting consisted mainly in alterations to the content. A Czech quality model called Framework for a school's self-evaluation (Kekule 2012) was used and there was no emphasis on gender mainstreaming. A few organizational adjustments were made in a) the composition of teams (groups of 2 to 5 evaluators were always made to work in one particular school, not multiple institutions) and b) work with the evaluation report (schools offered existing self-evaluation reports to the evaluators, who could then base their evaluation on them). As a consequence of these alterations, participating schools were offered slightly different documents/forms from those used for Gutknecht-Gmeiner's (2007) peer review.

In the project *Road to Quality*, peer review was understood as evaluation of a school by a team from another school. It was carried out as a mutual and systematic assessment of various activities (management, human resources, conditions for education, etc.) by a group of colleagues (e.g. headteacher, deputy headteacher, teacher, educator) from another school or various schools of the same type. It was always up to the headteachers of participating schools to decide about the composition of these teams. Peer review is based on previous self-evaluation for which every school has applied its own original approach. To a certain extent, peer review complements the self-evaluation processes in schools as seen by colleagues from other schools. Sometimes it helps the school evaluated to clarify the view of itself. Therefore, its nature is formative (developmental) rather than summative. Peer review is focusing on the development of the organization and improvement of professionalism rather than accountability and checking.

Thirty-two schools were involved in the activities of the peer review project. Originally, these activities were not planned as reciprocal, but as schools asked for this during the initial information meetings, reciprocity of 24 schools was arranged. This means that these schools experienced both roles, evaluated and evaluator. The remaining eight schools participated in one role only, either evaluated or evaluator.

The actual peer review process consisted of four stages:

First, schools were introduced in detail to the process of peer review they would undergo.

The lecturer[9] explained to the teams what they would do, how they would evaluate and how their own school would be evaluated. After the information meeting of school representatives, the time schedule of peer review was agreed on by the school evaluated and the evaluators. The team from the school evaluated agreed with the evaluators on the scope of items for evaluation. The evaluators received the self-evaluation report and other relevant documents (e.g. the school's annual report and/or School Education Programme) from representatives of the school under evaluation. Based on this information, they prepared for the visit, specified the criteria for evaluation, chose suitable methods and worded the questions for interviews and the questionnaire. The evaluators introduced the school evaluated to what they intended to focus on during the visit and what the school should prepare. The evaluators and the representatives of the school agreed in detail on the organization of the visit.

Secondly, during the visit, data collection was carried out by means of the methods the evaluators had prepared. School leaders, teachers and pupils were interviewed and teaching was observed. Sometimes it happened that the team of the school influenced the process of data collection by deciding who can or not be interviewed or visited by the evaluators. At the end of the visit, the group of evaluators gave the school preliminary feedback based on the study of documents and how the visit had unfolded. This feedback did not contain evaluative judgements. Stress was put on descriptive feedback and appraisal of what the evaluators had liked.

Thirdly, after the visit, the evaluators wrote a first version of the evaluation report. The report was then submitted to the school for comment by which those evaluated could raise objections to the findings of the evaluators. Comments were edited by the evaluators to achieve its final wording.

Fourthly, the task of the school under evaluation was to try and transform the results and recommendations into specific activities for improvement, and to plan and perform them. It was up to each school to decide which suggestions would be used or not for its own development. This was the link between the processes of evaluation and change. One problem of the voluntariness of the implementation of proposals for improvement was that schools could neglect the recommendations. This is also what makes peer review different from CSI evaluation; schools are obliged to implement the suggestions/recommendations for improvement before the following inspection visit or explain why they did not do so.

The participating schools were supported in terms of methodology. Throughout the process, including the visit, a facilitator was ready to provide assistance to both teams. This was an external specialist, remunerated for supervising the process of peer review. The role of the facilitators was played by professionals who were experienced in lecturing, studied the methodology for peer review as adapted from Gutknecht-Gmeiner (2007) for Czech conditions, and met the leaders of the *Road to Quality* in order to clarify uncertainties. Facilitator provided the school with professional support during the preparation stage, was at both teams' disposal, and acted as an intermediary during the visit and as a consultant for the elaboration of the report.

---

[9] These were two female lecturers provenient from the Faculty of Education, Charles University, Prague.

When specifying that peer review is a continuation to self-evaluation (see Fig. 7.1), it is necessary to bear in mind that (before peer review) schools carried out self-evaluation in a variety of ways; their experience differed, and their self-evaluation reports differed in both content and form. Every school focused their self-evaluation on different aspects and attached varied importance to it, which was later manifested by varied forms of reports. Self-evaluation was influenced by a governmental regulation (Regulation 15/2005)[10] that had stipulated six areas of quality: support from school to pupils; cooperation with parents; impact of relations in the school, pupils, parents and other persons; school leadership and management; human resource work, in-service teacher training; and content and course of education. Also, each school worked out these areas in different ways (see left section of Fig. 7.1 on self-evaluation) and used different evaluation tools. Based on self-evaluation, each school then worked out a plan for improvement and then created its own model of quality assurance in order to meet the legal requirement to conduct self-evaluation (Regulation 15/2005). Consequently, the school submitted to the evaluators a self-evaluation report that corresponded with its own model of quality

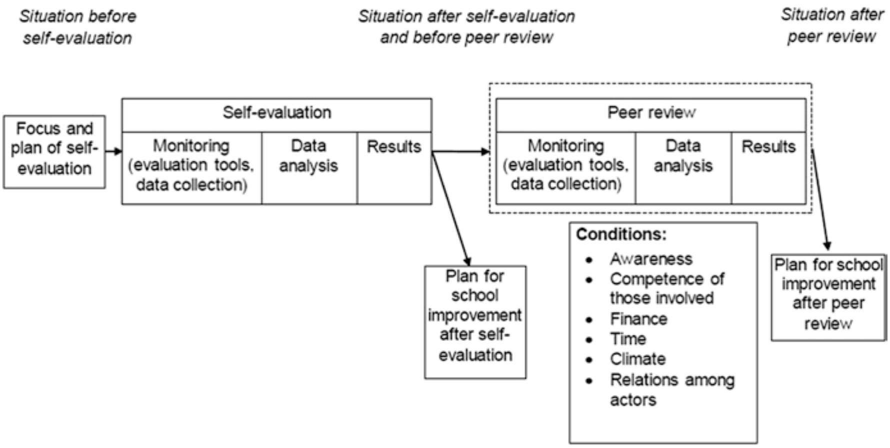

Situation before self-evaluation
Situation after self-evaluation and before peer review
Situation after peer review
Focus and plan of self-evaluation
Self-evaluation
Monitoring (evaluation tools, data collection)
Data analysis
Results
Peer review
Monitoring (evaluation tools, data collection)
Data analysis
Results
Plan for school improvement after self-evaluation
Conditions:
Awareness
Competence of those involved
Finance
Time
Climate
Relations among actors
Plan for school improvement after peer review

Fig. 7.1 Peer review as a continuation of the process of self-evaluation

assurance. During peer review, the evaluators used evaluation methods to collect data (for instance, they were preparing the questions for interviews and the questionnaire for pupils, specifying the items for observation), analyse them and make conclusions (right section of Fig. 7.1). The quality of the usage of evaluation methods was dependent on the experience of the evaluators. No framework of methods nor a model of quality were set up by the project team for peer review. Then, based on these conclusions, they provided the host school with suggestions for improvement, i.e. a plan for school improvement as a consequence of peer review. The content of this plan was different from the plan proposed by the school after self-evaluation. Based on the experience from the project *Road to Quality* and statements of the facilitators, we found out that 28 peer review sessions, different from each other, were carried out. In general, the course of peer review depended on the knowledge of appropriate procedures, competences for evaluation, interpersonal relations and accessibility of resources such as time and finance. It is likely that these differences were caused by different contexts of schools and the different experience they had with self-evaluation.

---

[10] It was during the project *Road to Quality* (2009–2012) when the criterional framework of the CSI changed (see Michek 2013, pp. 54–56) and was different from the quality areas stipulated by Regulation 15/2005.

7.5 Experience Gained from the Project: Research Methods, Objectives, Focus and Data Collection

The attitudes of school representatives to peer evaluation were investigated as part of the project *Road to Quality* (Chvál et al. 2012). This project took place from May 2009 to August 2012 and consisted of a variety of activities; here we describe those related to peer review only. A questionnaire survey was carried out in summer 2009 to learn about schools' interest in getting involved in the project; the schools addressed were offered the opportunity to try peer review in 2010; these schools participated in initial meetings as educational actions in autumn of that year; and peer review took place in the schools involved in winter and spring 2011. The data and findings as we present them are influenced by the willingness of school representatives (headteachers and deputy headteachers), who answered the 2009 questionnaire indicating interest in peer review, to take part in the initial meeting and involve their schools in peer review (2010–2011).

The *questionnaire survey* that was carried out in summer 2009 gave a good general picture on the state of involvement and understanding of SE and peer review and guided our data collection for phase 2, below. All schools of regional education (public-funded, private, church) in the Czech Republic (special nursery schools, basic schools, basic schools of art, secondary schools and language schools) were addressed. Headteachers of 531 schools responded to the questionnaire, which represents a return rate of approximately 8%. The questionnaire had three main objectives: (a) to find out what experience, needs and opinions the headteachers relate to self-evaluation; (b) to inform schools about activities they can join; (c) to enable schools to start these activities. Besides this, primary data made it possible to determine which evaluation activities had taken place in schools interested in peer evaluation (21% of schools addressed) and in those uninvolved (the remaining 79%). Some results of this survey have already been published (Chvál et al. 2010). These results are presented below.

In 2010 the project leaders contacted the schools that had expressed interest in peer review in the questionnaire and representatives of these schools were invited to information meetings. These meetings took place in June and November 2010; there were six altogether, attended by a total of 207 representatives of 46 schools (32 basic schools, 3 secondary schools of general education [gymnázium] and 11 secondary technical schools). The other schools did not follow up their interest owing to more urgent priorities and/or because their staff perceived participation in peer review as an organizational burden.

32 schools participated in peer review (26 basic schools, 2 secondary schools of general education and 4 technical schools). The representatives of the remaining 14 schools realized during the initial meeting what was expected of them, and withdrew. Twenty-eight activities were organized within the process; these were attended by 128 employees (43 participated once and 85 twice, the latter in both roles, as evaluated and evaluator). Of these participants, 54% were school leaders (headteachers, deputy headteachers), 16% were educators entrusted with duties in addition to their teaching (boarding house administrator, technical training manager, career guide, school education programme coordinator, ICT coordinator) and 30% comprised the other teachers.

7.6 Phase 2 of the Data Collection

Survey. At the end of the visit, lasting one and a half days, the evaluators and the participants from participating schools were asked to take part in a survey. Between October 2010 and May 2011, 181 respondents (85% of all participants) answered the survey. They were headteachers (23%), deputy headteachers (27%) and teachers (47%); 3% of them did not specify their function in the school). The composition of the group of respondents reflected the representation of all participants of peer review. The opening part of the survey consisted of questions aimed at possible continuation of peer review after the project Road to Quality and its assessment (fulfilment of expectations), applicability of experience/knowledge from peer review in further practice and organizational arrangement of the visit.

Facilitators' reports. The facilitators provided the leaders of the project Road to Quality by facilitators' reports with information on the course of the event, its benefits, barriers and difficulties in implementation, support provided for schools during peer review and the needs of facilitators as specialists assisting in evaluation actions. Twenty-seven record sheets were filled out by five facilitators. The role of the facilitator was played by two headteachers (both had been in office for more than 5 years), two lecturers of in-service teacher training (both had lectured for more than 5 years) and an academic who had worked over 10 years at a faculty of education.

Case studies. Of the 32 participating schools, four technical schools (providers of vocational education) were invited to be subjects of a more detailed research survey and produce case studies intended to analyse observations made during peer evaluation, semi-structured interviews and study of the documents submitted. These schools, of which we had no previous experience, carried out reciprocal peer evaluation. They were typically active in one way or another, be it willingness to participate in projects or be involved in foreign partnerships.

Observation. One of the researchers participated as an observer in two initial meetings at technical schools and took field notes on all relevant facts concerning motivation, also recording the remarks of peer review participants.

Semi-structured interviews with peer review participants from 4 technical schools. Four to two weeks prior to the evaluators' visits, we conducted semi-structured interviews with representatives of the four technical schools involved. The timing was chosen because teams from the schools about to be evaluated were already communicating with those ready to evaluate them, preparing intensively for the visit and, in relation to it, demonstrating expectations that were the point of research interest. At each school, an interview was carried out with the headteacher, while another was conducted with a member of the peer review team (deputy headteacher or teacher). The aim of these interviews was to learn more about the motivation of the schools in the peer review programme and their expectations of it. Another objective was to learn more about the competences of the participants and the position of peer review in the context of other evaluation activities at the school, particularly evaluation performed by other external subjects. One of the topics of these interviews was comparison of various aspects of self-evaluation with those of peer review.

One month after the evaluators' visits, more semi-structured interviews were conducted with representatives of the four technical schools in order to… The period of 1 month was chosen deliberately, as by then immediate impressions had faded, the assessment of the evaluation had been put into the context of other school activities and people at the school had begun to reflect on the recommendations of the evaluation. Interviews were held with the headteacher, one member of the team involved in peer review and one of the teachers who had not been a member of the team.

7.7 Results of Phase 2 Data Analysis

This section is a summary of the main results of our investigations into the implementation of peer review. Below is a discussion on the notion the representatives of participating schools had about peer review, how peer review contributed to the development of their schools and whether they would be interested in such an activity again in the future.

1. *Expectations of schools willing to carry out peer review*

The results presented here are based on findings from observation of initial meetings in autumn 2010 and interviews with representatives of four technical schools (both before peer review and after). The representatives mentioned the following expectations: gain of recommendations on how to self-evaluate (adopt the methodology, shape the structure of the self-evaluation report, learn to use new evaluation tools and so on), opportunity to perform evaluation in areas that have not been self-evaluated and learn partial procedures of evaluation in order to use them for the assessment of self-evaluation (meta-evaluation). One teacher expected she would learn to evaluate in terms of guessing promptly what is of key importance…

The opportunity to see the school from the outside was related to the following expectations: stimuli for school development, inspiration from and use of experience of others, opportunity to see a different school and colleagues, find out about views on the school held by colleagues from the outside, gain assessment of drawbacks and strong points of the school in specific areas, be compared to another school and discuss specific topics with colleagues from another school. An interesting comment was made by a deputy headteacher responsible for vocational training: "…it's that it takes the blinkers from my eyes, removes the operational blindness. There is this operational blindness, definitely". Specific expectations were mentioned, such as strengthening the control system at the school and starting new cooperation with another entity, as well as pragmatic ones. As another deputy headteacher said: "it seemed interesting to me, advantageous in terms of economics".

Also in relation to expectations, school representatives mentioned fears such as misgivings about the ability to evaluate another school "I suppose that at first the other school will have problems with it, because we're not trained to evaluate things" (a headteacher). In spite of this, they expressed readiness to cope with these burdens: "…but we'll manage one way or another" (a headteacher).

2. *Evaluation actions and work with data in peer review are beneficial for school development*

This proposition is based on findings gained from surveys, facilitators' reports and interviews with peer review participants in technical schools.

Respondents of surveys stated that peer reviews had fulfilled their expectations to a large extent. Minor differences appeared between basic and technical schools: 94.5% respondents from basic schools said their expectations were fulfilled, while in technical schools the share of respondents satisfied was 76.3%. A large majority of respondents from basic schools (94.5%) could imagine how they would use experience/knowledge from peer review in practice; in technical schools, such opinion was expressed by over three quarters of headteachers in this sector (78.9%). Specifically (see Table 7.1), this was mostly the case with the use of inspiration (incentive, idea, suggestion for school improvement), feedback, specific benefit and exchange of experience. The remaining quarter mentioned benefits such as the opportunity for comparison with another school, gaining an unbiased perspective of the school, the development of evaluation proficiency, insight into another place, the

Table 7.1 Beneficial for school development from peer review

Survey (N = 181)		Facilitators’ reports (N = 27)	
Case with the use of inspiration	41.2%	Exchange of experience and inspiration	21 occurrences
Exchange of experience	9.2%		
Feedback	14.5%	Feedback from schools of identical specialisation	7 occurrences
Specific benefit	9.2%	Specific recommendations for school development	12 occurrences
Opportunity for comparison with another school, gaining an unbiased perspective of the school, the development of evaluation proficiency, insight into another place, the opportunity to make a social contribution and motivation for work	25.9%	Encouragement, acknowledgement, words of support for further work	5 occurrences
		Development of ability to evaluate	5 occurrences
		Agreement on further cooperation	3 occurrences

opportunity to make a social contribution (e.g. acquisition of contacts at the partner school) and motivation for work.

Facilitators’ reports are completed by findings of the surveys (see Table 7.1). They explicitly highlighted the following features of peer review as beneficial for participating schools: (a) exchange of experience and inspiration; (b) specific recommendations for school development; (c) feedback from schools of identical specialisation; (d) encouragement, acknowledgement, words of support for further work; (e) development of ability to evaluate; (f) agreement on further cooperation.

Staff at participating schools used peer review to develop their evaluation skills. Representatives of evaluated schools usually acknowledged the importance of self-evaluation. The evaluators put value on how their abilities to evaluate were improved by methods like wording descriptive feedback to peers, asking questions, conduct of controlled structured interviews, creation and evaluation of questionnaires and relevant observation in classes of colleagues whom they had not known before. During the interviews after peer review, the participants often identified benefits for both themselves, i.e. individually, and for the school, institutionally. At the individual level, there was recognition of shared practices in other institutions by their peers (professional horizons) and these new personal contacts made further cooperation possible, e.g. telephone consultations in order to solve various problems or participation in competitions and other events. Peer reviews also provided information on strong points at the host schools that the evaluated teachers were not aware of or had considered standard.

The benefits for the school included confirmation of strong points and feedback on the shortcomings of leaders and other staff, information on which less obvious factors were nevertheless having a positive effect on the school, and the very welcome introduction to a partner school with whom they would be able to cooperate in the future.

> “We sat down after the event and talked about what it meant to us. It was just great. I’ve not found anything negative, there is nothing but positive impact. We’ve gone for it expecting a friend’s critical view, which would be a mirror that shows something to us. I think it was exactly that” (a teacher).

> “Meetings of schools are more meaningful than long-term trainings” (a teacher).

3. *The participants consider peer review different from inspection visits by Czech School Inspectorate and evaluation by school establishing entities*

This proposition is based on interviews conducted with peer review participants in 4 technical schools both before and after peer review.

The respondents perceived peer review as a qualitatively different from practices of the Inspectorate; in particular acknowledging the collegial and friendly approach of the former: “…peer review was a thousand times more pleasant than inspection (a class teacher)”. They were not afraid of peer review and could open up as if they were facing an equal partner. They acknowledged that peer review looked for strong points and offered recommendations, incentives for improvement, advice and exchange of experience, while the attitude to CSI was ambivalent, ranging from negative to neutral and positive judgements. As for negative attitudes, they said inspection was stressful, causing worries and fears:

> “…one is in fact a bit worried about inspection, because it’s a superior body that gives you a smack if they don’t like something, while this was… in my opinion… I’d say it was on an equal level” (a teacher).

Teachers had experienced bothersome inspectors who judged their work based on random sitting-in on classes. Another negative feature was that the Inspectorate’s focus was on faults and drawbacks: “…they simply have to find something wrong, otherwise it’s not real inspection” (a headteacher). As a rather neutral assessment, they often mentioned that the inspectors usually check the school’s documents, formal and organizational matters, regulations and compliance with legislation. They said inspectors assessed the state of education provided by the school but did not offer incentives: “…And I don’t think they can move me anywhere. It’s up to me to push this section forward” (deputy headteacher for vocational training). As to positive attitudes, the respondents said that the Inspectorate’s approach to the school was correct and inspectors were changing for the better, now evaluating strengths as well. Sporadically, inspectors were praised for having given advice: “…Cooperation is good there. I’ve got a feeling they give us a hand. So it’s rather consultation and advice on how to implement school the education programme in particular subjects” (a headteacher).

As to the municipalities (founders of public schools), schools do not expect incentives from them and, sometimes, consider their activities negative: “…They need to save money. There’s no other reason for any of their actions” (a headteacher).

4. *Participating schools are interested in further peer review activities with partner schools*

This proposition is based on surveys and interviews with headteachers of the four technical schools involved. Most survey respondents (87.3%) expressed interest in continuous partnership after the project. Positive attitudes to peer review are documented by the fact that a half of the respondents (50.3%) would like to have feedback from a partner school. Yet, there are differences between school types: 58.6% respondents in basic schools and 34.2% in technical schools want feedback from another school. Finally, only a quarter (28.2%) suppose that some school is willing to provide feedback for the school they work in.

Although the headteachers of participating schools commented on peer review as helpful, they were not sure about possible follow-up peer review with another school. It would be easier for them to decide: "…if there was an entity which could be directly addressed by schools. It would be definitely a benefit for us. We would be glad to use it. It was certainly positive for us" (a headteacher). When looking for a partner school for peer review they would focus on schools in the same fields of education but not competing. The headteachers pronounced that peer review would continue in that they keep in touch with the partner school, develop mutual relations, visit each other, talk on the telephone or write emails. The headteachers did not intend to carry out more permanent cycles of peer review in development with local partners schools if they were not organized in the form of a project.

7.8 Conclusion

Evaluation among colleagues who work in the same setting was not introduced in Czech schooling only in the past two decades. Nor is there anything new about the exchange of experience among teachers and headteachers. Various associations, networks, links and forms of cooperation were established many decades ago, as evidenced by some significant examples above. From the viewpoint of systematic management of and support for peer review, the project *Road to Quality* was a significant contribution. In this article we endeavored to describe the course of peer review within this project and summarize the main results of research surveys accompanying the preparation and course of peer review. We are convinced that some results are generalizable for peer review in other cultural contexts.

Activities aimed at the exchange of experience between schools are still under way in the Czech Republic. They are supported by finance from the European Social Fund as well as private foundations (e.g. The Kellner Family Foundation and *Učitel naživo,* see https://www.kellnerfoundation.cz/en/helping-schools-succeed, https://www.ucitelnazivo.cz/en). These follow-up projects are building on the lessons learned from the project *Road to Quality* (Štybnarová et al. 2012) and involve people who acted as facilitators in this project.

It turned out that peer review is a relatively difficult form of evaluation, demanding a considerable level of readiness and continuous support. On the other hand, those who participated have judged peer review very positively and documented indisputably its benefit for school development (cf. Michek et al. 2013). However, there is still a long way to a more extensive implementation of peer review in the Czech Republic. Schools involved in the project Road to Quality and the follow-up projects still account for a small minority. In order for peer review to be purposefully used by a higher percentage of schools it is necessary to arrange for financial, institutional and organizational support on both regional and national levels.

References

Auerswald, F. (1965). *Pedagogická tvořivost a pedagogická čtení* [Pedagogic creativity and pedagogic readings]. Praha: Státní pedagogické nakladatelství.

Auerswald, F., Hrabal, V., Chlup, O., Kníže, G., & Kočková, M. (1957). *Pedagogická čtení* [Pedagogic readings]. Praha: Státní pedagogické nakladatelství.

Bezdíček, J. (1938). Hospitace [Sitting-in on classes]. In O. Chlup, J. Kubálek, and J. Uher (Eds.), *Pedagogická encyklopedie* [Pedagogic encyclopedia] (pp. 479–480). Praha: Novina.

Chvál, M., & Starý, K. (2008). Autoevaluace škol v reflexi poskytované odborné podpory [School self-evaluation in reflection of specialist support provided]. *Orbis Scholae, 2*(3), 53–73.

Chvál, M., Procházková, L., & Černý, K. (2010). Postoje k autoevaluaci ředitelů základních a středních škol v ČR [Attitudes of Czech basic and secondary school headteachers to self-evaluation]. *Orbis Scholae, 4*(1), 91–110.

Chvál, M., Černý, K., Košťálová, H., Mechlová, E., Michek, S., Němec, J., Pol, M., Poláchová Vašťatková, J., Procházková, L., Shánilová, I., & Štybnarová J. (2012). *Školy na cestě ke kvalitě: Systém podpory autoevaluace škol v ČR* [Schools on the road to quality: The system of support for school self-evaluation in the Czech Republic]. Praha: NÚV.

Fundamentals of a Common Quality Assurance Framework (CQAF) for VET in Europe. (2007). Luxembourg: Office for Official Publications of the European Communities, 2007.

Gutknecht-Gmeiner, M. (Ed.). (2007). *European peer review manual for initial VET*. Vienna: öibf.

Kekule, M. (2012). *Rámec pro vlastní hodnocení školy. Metodický průvodce* [Framework for self-evaluation; a methodological guide]. Praha: NÚV.

Koncer, F. (1929). Hospitace [Sitting-in on classes]. *Školské reformy, 11*(3), 91.

Michek, S. (2008). Dva přístupy k externí evaluaci odborných škol [Two approaches to external evaluation of vocational schools]. *Orbis Scholae, 2*(3), 75–86.

Michek, S. (2013). *Externí evaluace odborných škol realizovaná partnery na stejné úrovni*. [External evaluation of vocational schools implemented by partners on the same level] Praha: Univerzita Karlova v Praze. Disertační práce.

Michek, S., Pol, M., & Chvál, M. (2013). Peer review in school self-evaluation: Cultivating skills in data use. *The New Educational Review, 34*(4), 30–40.

Národní program rozvoje vzdělávání v ČR: Bílá kniha. (2001). *National Education Development Programme in the Czech Republic; White Book*. Praha: UIV Tauris.

Pech, O. (1937). *Příruční slovník pedagogický* [Reference dictionary of pedagogics]. Praha: Nákladem Česko-slovenské grafické Unie.

Pedagogická čtení 1956–1957 [Pedagogic readings 1956–1957]. (1957). Hradec Králové: Krajský ústav pro další vzdělávání učitelů v Hradci Králové.

Peutelschmiedová, A. (2011). *Pedagogické návraty*. Retrieved from http://clanky.rvp.cz/clanek/c/S/14209/pedagogicke-navraty.html/

SSVS. (2011). *EVOS – garance kvality* [Programme for quality evaluation in vocational schools — Guarantee of quality]. Praha: SSVS.

Štybnarová, J. et al. (2012). *Vzájemné učení škol v oblasti autoevaluace* [Mutual learning of schools in self-evaluation]. Praha: NÚV.

Votava, A. (1928). Hospitace [Sitting-in on classes]. *Školské reformy, 10*(7), 221.

Vyhláška, kterou se stanoví náležitosti dlouhodobých záměrů, výročních zpráv a vlastního hodnocení školy. Vyhláška č. 15/2005 Sb. ze dne 11. ledna 2005 ve znění z roku 2009. [Regulation 15/2005 of the Czech Republic].

Zákon o předškolním, základním, středním, vyšším odborném a jiném vzdělávání (školský zákon). Zákon č. 536/2004 Sb. ze dne 10. listopadu 2004 [Education Act 536/2004 of the Czech Republic].

Stanislav Michek has worked at the Faculty of Education, University of Hradec Králové, as assistant professor since 2012. His professional interests are divided between school management, teachers' professional development and educational evaluation, recently focusing more intensely on educational evaluation. Previously, he worked at the National Institute for Education for eight years, where he devoted himself to quality assurance in vocational education and training (VET) and helped lead the national project Road to Quality (2009–2012) as Deputy Chief Manager. In 2006–2009, he was a member of the ENQA-VET (*European Network for Quality Assurance in Vocational Education and Training*). Currently, he is also board member of ČAPV (*Czech Educational Research Association*).

Martin Chvál has worked at the Faculty of Education, Charles University, as researcher since 2007. His work focuses on self-evaluation, external school evaluation and student assessment. He was the Chief Manager of the project Road to Quality (2009–2012), focused on the support of school self-evaluation. He has been engaged in the field of education at both national and local level: as a member of the Independent Expert Committee for the Czech School Inspectorate and as advisory body member of the Ministry of Education's work to determine the validity of entrance exams and state school-leaving exams. He is also deputy chairman of the *Czech Educational Research Association* and a member of the executive editorial board of the journal *Pedagogika*, where he was chief executive (2013–2018).

Milan Pol has worked at the Faculty of Arts, Masaryk University, as Professor of Education and, since 2014, Dean of Faculty. His professional interests are divided between school leadership/management and educational evaluation, recently focusing more intensely on evaluation in higher education settings. Among other subjects, he has been involved in research in school culture, school governance, organizational learning in schools and life/professional careers of school leaders. He is author and co-author of a variety of home and foreign publications and has been editor-in-chief of the journal *Studia paedagogica* since 2009. Currently, he is also board member of ENIRDELM (European Network for Improving Research and Development in Educational Leadership and Management).

Part V

Peer Review Within School Improvement Partnerships

Chapter 8
The Development of a System Model of Peer Review and School Improvement: Challenge Partners

George Berwick

Abstract In this chapter, the use made of a peer review to initiate knowledge transfer and creation across a large network of schools is described. The peer review process is carried out by senior staff from other schools in the network. The review's focus is the effectiveness of the school's self-evaluation against a framework predominately reliant on Ofsted, the accuracy of its evaluation of teaching and learning, and the resulting validity of the school's action plan. In addition, the school being reviewed can identify areas it considers to be excellent which can then be accredited by its peers. A description is then given of the various approaches used to ensure the rigour of the process, the theoretical underpinning and the underlying design principles. Finally, the impact of the process on the school being reviewed and those undertaking the review is discussed.

8.1 Introduction

This chapter a description of the Challenge Partners peer review process administered annually across 400-plus schools, in all phases and a wide geographical spread in England. This is followed by a description of how rigour is ensured, the theoretical underpinning of the process and the incorporation of Challenge Partners' five design principles which have merged from its theory of action. The other linked activities are then noted. The chapter concludes with results and conclusions.

G. Berwick (✉)
georgeberwick.com, London, UK
e-mail: info@georgeberwick.com

© Springer Nature Switzerland AG 2020

D. Godfrey (ed.), *School Peer Review for Educational Improvement and Accountability*, Accountability and Educational Improvement,
https://doi.org/10.1007/978-3-030-48130-8_8

8.2 The Growth of the Partnership

Challenge Partners was borne out of the London Challenge school improvement programme. Running from 2002 to 2011, this was designed to raise the performance of schools in London through a series of connected initiatives. One of these was the creation of Teaching Schools. Loosely based on the concept of a Teaching Hospital, these schools needed to be outstanding in their context and to have the capacity to support underperforming schools. Collectively they would form a network of schools across the city. They would work to transform low attaining schools whilst simultaneously improving their own standards. Initially there were four across London, three of which later formed part of Challenge Partners. The emerging lesson of this work was that strategic and systematic school-to-school work had a major impact on standards in *all* schools. The fundamentals of this approach were an unreserved commitment to improving the performance of all students, especially those from disadvantaged backgrounds, through open accountability, effective knowledge management and building trust. The schools providing the support were high performing and they were given time and assistance to build the capacity to share their knowledge with their less successful peers. These schools became the Teaching Schools and those they supported formed a hub that was often joined by other schools on a voluntary basis. Their purpose was to learn from – and with – each other.

By the end of the Challenge in 2011, a theory of action had emerged from this effective school-to-school work (Berwick 2011, 2013, 2015). Within this, the concept of a Teaching School working with other schools to form a hub and for these hubs to be connected into a regional framework was the proven organisational framework. Forty four teaching schools had been designated by then, the majority in London but also in the other Challenge regions of Manchester and the West Midlands. This gave the confidence for the then Coalition Government to further approve the designation of hundreds of new Teaching schools, starting with 100 in 2011 (HM Government 2010). As of June 2018, there were 835 Teaching Schools across England.[1] Over this time, the Challenge Partners network also grew, both in number and in geographical spread, and included a higher proportion of primary schools. Between 2011 and 2018 the average size of hubs doubled from seven to 14 schools (Challenge Partners 2014b, 2015b, 2019). Some hubs have grown and then divided and one has done this twice. Others have continued to grow without dividing – the largest contains 35 schools. As more Multi-Academy Trusts have formed within the Partnership or have joined, the growth of the hub has been controlled by expansion rather than the voluntary growth which was the dominant driver in the early years. Challenge Partners itself has also evolved to include other elements of school improvement and professional development. These include a subject directory of quality assured examples of excellent practice and staff with capacity to share this practice and leadership development days to showcase a school's effective management of a new initiative or a new innovation it has developed.

---

[1] https://www.gov.uk/government/publications/teaching-schools-and-system-leadership-monthly-report/teaching-schools-and-system-leadership-june-2018

Challenge Partners initially consisted of 71 mainly London-based secondary schools formed into 12 hubs (Challenge Partners 2013). Growth across the period of the case study has been continuous. By March 2018 the Network of Excellence consisted of 422 schools networked in 33 hubs (Challenge Partners 2019). In the partnership as a whole, the schools were spread across all of the country apart from the North East of England. In March 2018, a total of 290,000 pupils were in the schools which formed Challenge Partners.

8.3 The Quality Assurance Review

8.3.1 The Process

The members of Challenge Partners openly challenge each other to provide a quality education for all of their students. This is built around a professionally-led peer review which focuses upon school improvement, with an emphasis upon the quality of leadership and teaching and learning (Challenge Partners 2014a, 2015a). The Quality Assurance Review (QAR) process:

- Validates the school's improvement action plan.
- Accredits areas of strength which can be shared within the hubs and across the partnership.
- Identifies areas for development, bringing key challenges to schools for the coming year.

There are six stages to the process. See Fig. 8.1:

1. **Scheduling**: Before the process starts for a school, the Challenge Partner central team will have worked with all of the schools and the hub managers to schedule both the reviews and the review teams. In addition, they will have ensured that the lead reviewers are selected, trained and deployed. During the period of the case study, this task has increased by more than 500 per cent. Not all schools have a QAR every year. They can opt out if they have an Ofsted and the majority in this case choose to do so. There are other local circumstances when schools decide not to participate but these are rare, and no school has gone more than a year without a review. The annual take-up rate has remained at around 90 per cent.
2. **Training:** Reviewers and the lead reviewers are trained in the process and protocols of the QAR. This usually occurs in June and July. This is non-negotiable.
3. **Pre-Review Analysis**: Prior to the visit, the review team determines, with input from the school, the timetable for the review. The team then analyses the school's performance data which is held in the public domain. The school being reviewed then provides its plan for improvement. If they choose, the schools being reviewed can provide additional information.

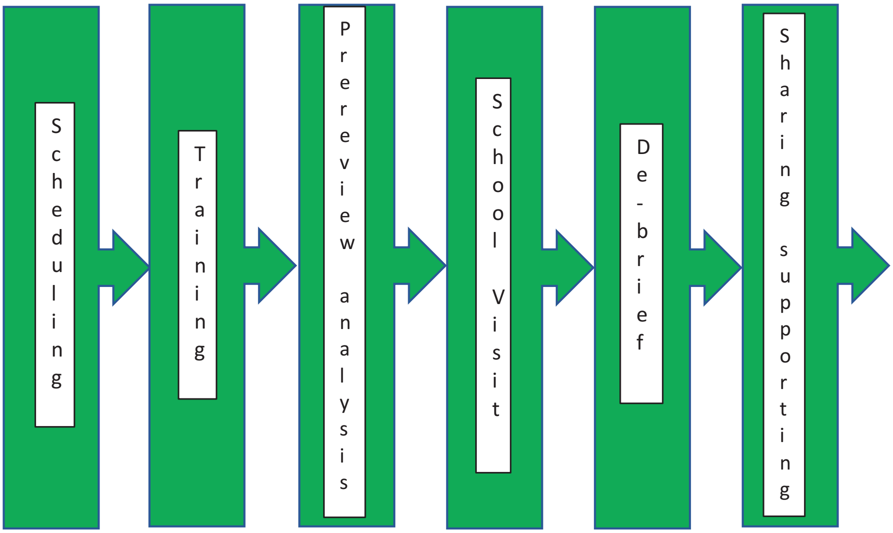


Fig. 8.1 The six stages of the quality assurance review process

4. **School Visit**: These take one and-a-half days. They include joint lesson observations and meetings with pupils and key staff. Visits are conducted by a review team of up to four senior leaders from other schools and led by an Ofsted-accredited inspector. The visits commence in September and initially run to April. However, with the increase in numbers, this has been extended into May and June.

During the review, the school can choose to identify an area of excellence to be validated by the review team. The review team will make its validation by referencing the appropriate Ofsted criteria, pupil performance information and its own peer judgement. If the team accepts this is an area of excellence, then the school details will be published in the subject directory. If the central team then considers it viable, from its knowledge of the demands schools are making for support in this area, it will ask the school to run a leadership development day or immersion day alongside a support service for other schools through the national brokering service.

1. **Debrief**: The review team produces a report. This is a collaborative process between all involved. There is an expectation that the school will share the report both internally and externally. Additionally, the reviewers receive feedback on their role, which they share with their own school as part of their Continuing Professional Development (CPD).

1. **Sharing the outcome with peers**: After the review, it is expected that the school being reviewed will incorporate the findings within its school improvement action plan. This might mean:

- Confirming or adjusting its perception of where they are now.
- Making adjustments to its self-evaluation of teaching and learning or their performance data.
- Confirming or adjusting their priorities for school improvement or introducing new ones.
- Identifying areas of knowledge it needs to acquire from elsewhere in the network.
- Building the capacity to share its strengths with other schools in its hub and nationally.

The school then shares the outcome with the other schools in its local hub. It does this to seek local support to resolve its needs and to offer support to others with areas that have been deemed outstanding by the review team. This is part of the collaborative process.

Those reviewing the school have an opportunity to develop their own skills at school evaluation. There is an expectation that this will form part of their personal development plan. In addition, they will take back to their school new knowledge about how they might improve and incorporate this into their school improvement action plan.

8.4 Ensuring Effective Collaboration

The QAR is a collaborative process. All observations and meetings include a member of the school team working alongside the reviewers, and evaluation of the findings is a collaborative effort between the review team and the school's senior leadership team. This approach enables honest and open conversations about where the school is and where it is going, to the benefit of all concerned. The dialogue that takes place is professionally focused and seeks a degree of consensus, but this is negotiable. The richness and the variety of the experience the reviewers bring to the debate is part of the value of the process. The expectation is that all involved will benefit.

This self-regulation extends across the process. For the majority of the activity within the QAR process, a set of non-negotiables has been identified. These are actions which over time have proved necessary to ensure that best practice is adhered to. If best practice has not been identified for a given activity, a degree of variance might have been established. Neither are legally binding for members, but it is the expectation that the action will be carried out within this agreed framework. An example of a set of non-negotiables would be that schools are expected to provide competent senior staff to act as reviewers, who will respect the privilege they have been given of visiting another school and have attended the reviewer training programme provided by the central team. It is their responsibility to the other schools in the network to do so. The schools being reviewed are expected to make adjustment to their school improvement action plan accordingly, and to share their outcomes with the other schools in their hub. The reviewers in turn are expected to share their new-found knowledge with their colleagues. All these actions rely upon self-regulation – which is an essential part of the reliability of the process – and collaborative working.

As part of developing a theory of action for the process, the non-negotiables and degrees of variance are constantly held under review. As we learn how to undertake the process more effectively from the wide range of feedback we receive, so we make adjustments accordingly. In the early days this was often on a month-to-month basis. As the process has become more refined, however, so these adjustments have been made annually to ensure they are part of the appropriate training programme and all the schools have been briefed.

Once the analysis was complete, each member of staff was required to identify an under-performing group of their own students and work collaboratively using the same sources of knowledge to construct an appropriate learning programme. They would then implement this programme in their schools and jointly analyse the outcome on a future visit.

8.5 Ensuring Rigour

The rigour of the review is provided during the process by the triangulation of the knowledge of school improvement from the practitioners – the school being reviewed and the reviewers – the lead reviewer and national standards encapsulated in the Ofsted framework for judging schools (OFSTED 2019) and the Department for Education in the measures it uses to assess school performance (Department for Education 2020): Fig. 8.2.

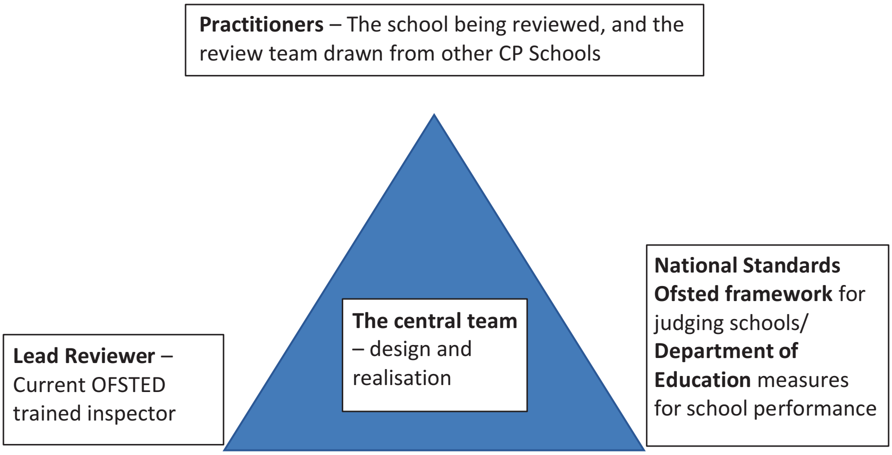

Practitioners – The school being reviewed, and the review team drawn from other CP Schools
The central team – design and realisation
Lead Reviewer – Current OFSTED trained inspector
National Standards Ofsted framework for judging schools/ **Department of Education** measures for school performance

Fig. 8.2 The four components of challenge partners QAR

8.6 The Practitioners

Challenge Partners does not select the schools which join it, nor the staff who take part in any of its activities. The hubs are run by Teaching Schools and they manage the selection procedure. Schools opt in. However, there is a local vetting process which makes a judgement about a school's willingness to learn with its peers and the capacity it has to do this. Once invited to join and on paying their annual fee, schools take part in a thorough induction process. The objective here is to ensure that schools do not view the QAR as a one-off but as part of a holistic approach to school improvement which will require them to make a considerable commitment beyond the joining fee. Schools see this as a cultural approach. Working with peers in this way is not the norm. There is an assumption that all schools in the network have something to learn.

All the practitioners involved have experience of Ofsted. For the majority this has occurred several times. A small but significant number have been or are trained Ofsted inspectors. They all therefore, from their varying experiences, understand the vagaries of the process and adjust their approach accordingly. They know that the Ofsted process can be contentious and guard against this with their colleagues without resorting to difficult questions. All are aware that this is not an Ofsted inspection. This is a peer evaluation of a school's performance and its resulting school improvement action plan with reference to national standards. Therefore, the comments made do not have a validity that could be used in any formal way. They are developmental and for guidance and only for the internal consumption of the school being reviewed.

The reviewers are drawn from a wide range of schools – geographically, socio-economically – performance and phase. By having a team drawn from at least two schools plus the lead reviewer, the review team's work is triangulated. They must also undertake the reviewers' training programme.

At the end of each review, the subject school and the review team evaluate the process together. This includes a judgement about the effectiveness of the lead reviewer. This is taken into account when lead reviews are evaluated and by the central design team which makes adjustments according to the design.

8.7 Ofsted and the Department for Education

Both the Ofsted Framework and the measures of the Department for Education (DfE) for student performance were modified during the period of the case study. For example, the DfE introduced new measures for secondary schools' performance such as progress 8 (Department for Education 2014a, 2016). So, in order to retain the rigour of the process, the reviewers and schools being reviewed had to remain current and make adjustments to their evaluations accordingly. These changes did on occasion take place within a school year. This was countered by having as many lead reviewers as possible with current Ofsted experience.

During the period of the case study, modifications were made to the ambition and criteria for Teaching Schools (Department for Education 2014b). In 2014, the DfE described them as, “Strong schools led by strong leaders who work with others to provide high-quality training, development and support to new and experienced school staff.” (ibid)

Using the term ‘strong’ rather than ‘outstanding’ in this statement broadened the eligibility range of schools and their leaders, from only those rated ‘outstanding’ by Ofsted to include those rated ‘good’.

At the same time, the criteria for designation, previously open to all schools in the country, was narrowed to those located in specific areas – the ‘cold spots’. These areas had been identified as having a record of poor student performance. The criteria were:

- *Evidence of providing support into published ‘cold spot’ areas and what plans exist to grow that support.*
- *Evidence of how a school has engaged in leading, managing and quality-assuring the training of teachers.*
- *Evidence of collaborating with other schools and partners, in the planning, design, development, delivery and evaluation of continuing professional and leadership development provision (beyond that expected of a normal school).*
- *Evidence of providing significant formal support over the previous three years (including details of the schools, the headteacher names, URNs and local authority districts) and an overview of that support and the impact it has had.*
- *Supporting narrative in relation to a school’s performance data. (ibid)*

This narrowing of the location of schools which were eligible to apply whilst reducing the performance criteria could in the longer term have an impact on the growth of Challenge Partners. However, the partnership was already represented in the majority of these ‘cold spots’ and there is no evidence to suggest these changes impacted the rigour of the process.

8.7.1 The Lead Reviewers

The process is managed and led by the lead reviewer. This individual not only requires the credibility associated with being an Ofsted-trained inspector but also the skills to facilitate the collaborative learning process. This requires a high level of social capital. The Ofsted process for many schools can be highly traumatic. Schools know that changes to their category can affect funding, status in the local community and the careers of the staff involved. To regulate this, lead reviewers are selected carefully, field tested, evaluated for their performance after every review and have their skills continually honed. We were fortunate that the senior lead reviewer throughout most of this period was Joanna Toulson, who was not only a well-respected Ofsted inspector but a highly successful headteacher. Her collaborative approach served as an inspiration for those who worked with her.

8.7.2 The Central Team

The central team provides rigour by ensuring that the design of the process represents clearly the Challenge Partner's theory of action and that the delivery of the process achieves high standards.

8.8 The Theoretical Underpinning

The school–to–school theory of action which has been developed over the past 20 years through collaborative action research provides the theoretical underpinning for the process. The theory provides a number of models, concepts and principles used to shape the action and non-negotiables and degrees of freedom to regulate it.

8.8.1 Tacit and Explicit Knowledge

Staff in schools work for the majority of the time using tacit knowledge. In order to develop their knowledge and especially to share it with others, they need to have periods of time when they can make it explicit. The design of the review process allows the two parties involved the opportunity to develop their tacit knowledge by making it explicit at different points in the process. See Fig. 8.3.

8.8.2 Stages of Learning Activity

In order for staff to learn effectively, they need to be provided with knowledge in three different forms:

(a) Pure information in a variety of mediums
(b) Experiential
(c) Coaching and mentoring.

Not all the forms are required and the key determining factors will be the complexity of the task to the learner and the social, emotional and political content of the knowledge. As a large proportion of the knowledge about schools is of the latter, activities involving (b) and (c) are essential to the design. Thus, the QAR includes a

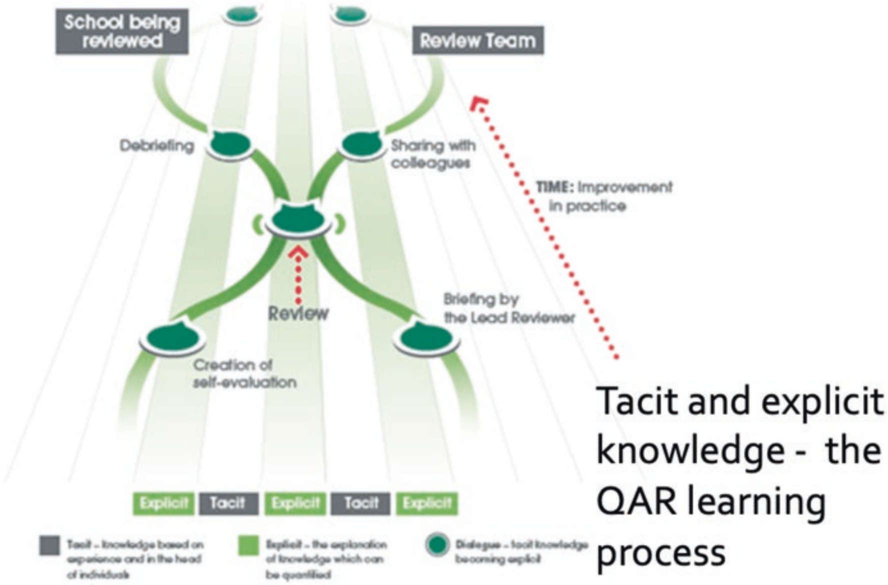

School being reviewed
Review Team
Debriefing
Sharing with colleagues
TIME: Improvement in practice
Review
Briefing by the Lead Reviewer
Creation of self-evaluation
Tacit and explicit knowledge - the QAR learning process
Explicit
Tacit
Explicit
Tacit
Explicit
Tacit – knowledge based on experience and in the head of individuals
Explicit – the explanation of knowledge which can be quantified
Dialogue – tacit knowledge becoming explicit

Fig. 8.3 The review process and the development of tacit and explicit knowledge (Challenge Partners 2014a)

visit to the school site, observation of lessons and talking to staff and students – experiential learning and the facilitation of collaborative learning by the lead review, which will include opportunities as necessary for indirect or direct coaching.

8.8.3 Using a Collaborative Learning Model

The manifestation in most schools of this process is in the presentation of a school improvement action plan. This plan is the result of the school asking the sequential questions:

a: What is it trying to achieve?
b: Where is it now?
c: Where should it go next?
and
d: How will it go from A to B?

In schools, the development of the school improvement action plan and its implementation follow a fairly regulated timeframe determined by the provision of resources in April or September and the publication of examination results in July and August. Less regulated is the visit of Ofsted, which is at random times of the year and over different yearly cycles depending upon the school's performance.

However, because of the large number of reviews taking place annually for a small but significant number of the Challenge Partners schools, the timing of the review does not synchronize with their cycle of school improvement action planning. This is a potential weakness in its design. In practice, the schools have accepted that because of the number of reviews being undertaken in some years, it is inevitable that they are out of line and have traded this off against their perceptions of the longer-term benefits of the process.

8.9 The Issue of Context

We have found that schools tend to learn more from peers operating under a different context than from similar ones. Context is the premier barrier to knowledge transfer, however, because difficult learning issues in education are often so personalised. The lenses of a different context draws out the underpinning generic issues. Thus, in the QAR, schools cannot select which phase or location they are to be sent to. This approach ensures that the schools access the totality of the knowledge in the partnership. It is also assumed that the schools have already learned what they require from their peers in their hub or are in the process of doing so.

8.10 'You Get Nowt for Learning'

Within the partnership, schools do not receive payment for their staff's learning. This is a principle we have all adopted rather than a proven rationale. We argue that we provide a paid service which is valued by our schools, many of which are rated outstanding. If they did not consider the product valuable, they would not pay for it. This is a rigour that we have embraced and often delineates our work from others.

In practice this means that the reviewers are not paid for their absence from their school or for the cover required. The school being reviewed is not paid for its staff time either. The only payment the schools make is to the central cost of running the QAR and the partnership. This payment is made annually and this means that schools can opt in or out on a yearly basis. The annual retention rate is around 80 per cent. The major reason for schools leaving is a change in the leadership of the school or group of schools. The brevity therefore of some schools' involvement in the partnership is at issue with our model of school improvement, which requires a high degree of continuity over several years.

8.11 The Importance of Leadership and Teaching-and-Learning

In the theory of action, the key components of school improvement identified are the quality of leadership and then the quality of teaching and learning (Berwick, 2011, 2013, 2015). Without effective leadership, schools do not function. This raises a key issue for the QAR. Even if the QAR validates a school's analysis of where it is now, where it should go next and how this should be achieved, if the school's leadership is not able to implement this, the required action identified is then worthless. The dilemma lies in how to regulate this in a self-regulating system. One way has been for the central team to identify schools already in or about to join the partnership which have been singled out by Ofsted as under-achieving. If this is the case, they then check with the teaching school leading the hub to ascertain if it has the capacity to implement its school improvement plan. If this is not the case, then a more forensic review will be suggested. Fortunately, this has seldom occurred.

8.12 Ensuring that the Knowledge Being Circulated Is Worthwhile

During the lifetime of the case study, we have been approached by a number of groups of schools which have wished to set up their own programme based on the Challenge Partners model. A number of these have been under-performing groups of schools. Our challenge to them has been, "why would you want to circulate the knowledge you have when the evidence is that it is not working?" Within the Network of Excellence (NoE) there is a larger proportion of the highest-performing schools than the national figure and this ensures that the knowledge we circulate is worth knowing.

8.13 The Five Principles

Challenge Partners was fortunate to be sponsored by the Social Business Trust. The charity draws on the resources of a number of major companies to support charities in the areas that they might struggle, such as marketing, finance and business management. In our case they provided us with expertise in product development from the international consultancy firm, Bain and Co. Their team worked with the trustees, senior partners and central teams to review all of its programmes. The outcome of this was twofold.

For the QAR, the central team provided the schools with detailed documentation that described the process being used and the expectations of each group involved, the reviewers, the team being reviewed, the hub manager, the central team and the lead reviewers. It used non-negotiables and degrees of variance to state the expectations of all involved. This was a live document which was updated as the process was continually refined.

For all of Challenge Partners' programmes, the central team published the five guiding design principles (Challenge Partners 2018). These would be used to shape new programmes and to evaluate existing ones. The objective of using these design principles was to achieve the aims of Challenge Partners. This meant improving the life chances of all the students in the collaborative of schools, especially the most disadvantaged.

The principles were:

Principle 1
Practitioner-led and engaged

School practitioners started Challenge Partners and should be actively leading the system, making decisions that drive pupil outcomes, with a strong voice in national policy. We seek to engage practitioners in our activities and believe the more engaged they are the more they, and ultimately the pupils, will benefit.

Principle 2
Share and improve practice through collaborative learning

We facilitate learning and effective practice sharing through collaboration to reduce the variability and drive the best individuals to grow new practice. Drawing upon effective practice, we share and create new knowledge to improve the quality of leadership and teaching. We invite and provide each other with both challenges and support to provide the best for our pupils.

Principle 3
A national network of local partnerships

We develop collaborative communities at school, local, national and international levels around Teaching School Alliances, multi-academy trusts and other existing partnerships, and so create a scalable model for national reform.

Principle 4
Access the best from within and beyond education

We look beyond ourselves and reach both into the wider education sector and into other sectors, including public and private sectors, for effective practice and support.

Principle 5
Create sustainable solutions

We seek to create sustainable school improvement programmes, and by pooling resources, schools get more out than they put in.

In 2017, these principles were used by the leadership team and trustees to evaluate the Peer Review. They used a RAG (Red, Amber, Green) rating. For the QAR they concluded that for the aim of the process, improving the performance of students, especially the most disadvantaged, the result would be amber (Challenge Partners 2018). They argued that though the schools in the network were achieving their aims, the link between the QAR and the outcomes was too tenuous to claim a high degree of causality and thus award a green. In their view there was clear evidence of relationship; schools in Challenge Partners undertook the QAR and their results improved. However, far too many other factors needed to be taken into account than just the QAR to claim causality. Also, for all the rigour and reliability built into the process, it was still voluntary and dependent upon self-regulation.

They accepted that in the design of the process, Challenge Partners had gained considerably from the insights provided by Bain as to how other industries ensured consistency of practice across a wide range of providers. However, at present there was no overt involvement beyond this.

8.14 Other Linked Activities

As has already been identified, the QAR is the part of a school's improvement process adopted by a Challenge Partners school based on its theory of action for effective school-to-school work. Resulting from the completion of the QAR, a number of activities arise at hub and national level.

8.14.1 Activities Provided by the Local Hub

If the centrally administered QAR presents a nationwide challenge for Challenge Partners schools, the majority of the support is provided at a local level by the hubs. The information contained in each school's QAR, augmented by that which is in the public domain and their knowledge of working collaboratively with the school, allows the members of the hub and especially the leadership to build up a picture of each school's capacity and requirements for their own and other schools' improvements. From this they establish a clear notion of what support is required and the local capacity to provide it. Over time, as with all activities in the NoE, this analysis has become more refined.

This is the start point for the collaborative school improvement programme at hub level (Berwick and John 2017). Using their intelligence about the schools in the hub, Challenge Partners pool resources and allocate them strategically. For example, they might with reference to the latest information provided by the Education Endowment Fund (EEF), decide to introduce a literacy programme to the schools or provide a training programme designed to eliminate poor quality teaching. The latter could result in some schools gaining a larger share of the resource than others. The schools accept this as part of working collaboratively. In most hubs, funds from the charity are provided to assist this. Schools often self-select each other to work together on common issues. These can result in permanent sub-networks, for example meetings between headteachers to discuss English or a numeracy teacher group. This is needs-driven and has often been identified as a result of the QAR.

To co-ordinate these school improvement activities, the schools meet regularly at a local level and the hub managers and headteachers leading each hub attend meetings with the central team each term to share best practice and determine the future direction of the charity. The successful role modelling of collaborative leadership by these two key practitioners is critical to the development of the hub. Our experience is that if they set the tone by being perceived by colleagues as trustworthy, open, empathetic, resilient and as having a determination to see every pupil within their hub receive the best education they can provide, then the hub prospers. If, however, the other schools perceive that they are only leading because they want to protect their own knowledge and control access to new ideas, then the hub stagnates or eventually the schools leave.

Also, critical to ensuring how a hub continues to succeed is how it selects and inducts new schools and new headteachers to work with the existing schools in the hub. It needs to know that the new recruits believe in working collaboratively and will do so. If they cannot work collaboratively in this way, they will view the QAR as an isolated activity and not as the annual start point for their collaborative journey of school improvements, challenged and supported by their peers. As a result, they treat it as a service-level agreement, often as a mock Ofsted and do the minimum to engage with other schools. These schools often leave after one year. However, due to the quality of the QAR, it is not a major reason for schools leaving Challenge Partners.

8.15 Activities Provided by the Central Team

Along with administering the QAR, the central team is involved in three other related activities. These are:

- Facilitating the national brokering service
- Administrating the subject directory
- Facilitating leadership development days

8.16 National Brokering Service

As has been previously stated, once the QAR is complete, schools are expected to share the report with the others in their hub. If the report identifies that they require support, then the hub is the first port of call. However, if the hub does not have the capacity to meet their request then they can contact the central team, which brokers support from across the network. Either the school requiring support or the hub manager would initiate this contact. The provision of this support is referred to as the National Brokering Service (Challenge Partners 2016).

Two areas that have received limited but constant demand for national support are for experienced department leaders with a track record of success to coach their peers in other schools, especially when a new person is in post, and in specialist subjects were there is a small pool of teachers locally.

8.17 The Subject Directory

The publication of a subject directory has fuelled this demand. It contains the names of those schools which have had an area of excellence designated during their QAR. The information in the directory allows schools to contact each other directly (Challenge Partners 2018).

8.18 Leadership Development Days

Another way the results of the QAR are used to share best practice across the network is through leadership development days and other forms of on-site visits offered by schools to allow their colleagues to experience best practice first hand (Challenge Partners 2018).

8.19 Results

In this section the overall results are presented. They are divided into three categories; internal, those generated by Challenge Partners and those external from other sources. The internal results start with the NoE performance against its four aims as reported in the 2018 Challenge Partners Annual Report (Challenge Partners 2019). These provide the context for the results of the QAR that follow. The external results are taken from an independent review of the QAR undertaken in 2015.

8.19.1 Internal

The results for Challenge Partners 2017–2018 (Challenge Partners 2019).

Aim 1
To improve pupils' examination results at a rate above the national average.

In this year the consistent pattern of performance and public examinations was affirmed. In primary and secondary attainment, students in Challenge Partner schools on average score above the national average. Progress in primary students at Key Stage 2 (Age 11–12) reading, writing and mathematics is above the national average and in secondary at Key Stage 4 (Age 15–16) the same pattern emerges with 20 per cent of students well above average compared to 12 per cent nationally. The gap in performance between disadvantaged students and other students is below the national average.

Aim 2

Enable all schools to improve at a rate above the national average (Table 8.1).

Sixty-five schools in Challenge Partners had an Ofsted inspection in 2017–2018, with 48 per cent of them improving their grade compared to 34 per cent nationally. This percentage increases to 52 when schools which have been in the NoE for at least three years are selected.

Numbers of schools in the Ofsted categories above vary from year to year, however there is a notable pattern of above-average transformation from one category to another in Challenge Partners schools.

The moral purpose of Challenge Partners is to bring in underperforming schools and support and challenge them in their transformation, thus resulting in the cyclical nature of participating schools at any given time. Aggregated results for participating schools will therefore reflect the changing capacity of existing schools in CP to support the underperformance of new arrivals.

Aim 3

Create more outstanding schools that reach the Teaching Schools criteria with national leaders in school-to-school work.

Over the period of Challenge Partners there has been an increase in the number of Teaching Schools, National Leaders in Education, and National Support Schools. Proportionately there have been more of these represented within CP than the national average. Currently within the NOE there are (Challenge Partners 2019):

- 38 Teaching Schools
- 31 NOE hubs
- 13 clusters of schools which deliver a Challenge Partners programme called Challenge the Gap. This is designed to reduce the gap between disadvantaged students and the mainstream.
- 11 delivery centres for a programme designed to improve the ability of teachers to educate students for whom English is a second language.

Table 8.1 Detailing Ofsted category percentages for English state-funded and Challenge Partner Schools

Ofsted category	National	Challenge partners
Outstanding	21%	26%
Good	65%	62%
Requires improvement	9%	9%
Inadequate	2%	2%
		Figures rounded

In 2017–2018

- 943 headteachers and senior staff were trained to act as reviewers with a number acting as mentors
- 111 schools had accredited areas of excellence
- 182 headteachers and senior staff attended leadership development days.

Aim 4

Develop a world-class, self-improving and sustainable system that contributes to national excellence.

Challenge Partners' headteachers, teachers and the growing number of CEOs of multi academy trusts (MATs) play leading roles in their regional and national organisations designed to lead the school system. This in England is currently a school-led system and therefore has considerable influence.

8.19.2 The Results for the QAR 2015–2018 (Challenge Partners 2016, 2017, 2018, 2019) (Table 8.2)

The table records that the majority of the schools choose annually to have a QAR. However, 10 per cent of the schools in the NoE still choose not to have a QAR. The main reason for this is that schools often opt out if they receive an Ofsted inspection in the same year. However, this is not always the case. For example, in 2017–2018, of 65 schools which had an Ofsted inspection, 25 elected to also have a QAR.

With on average 1,000 senior staff acting as reviewers annually and 400 schools receiving a review, the number of senior staff involved annually will be in the region of 2,000–3,000. The result of this is that the connectivity of the network is continually being refreshed and senior staff engagement in school improvement is annually assured.

Table 8.2 The data collected by the Central Team concerning the QAR 2015–2017

	2015–2016	2016–2017	2017–2018
% of schools having a QAR	88%	90%	89%
Number of reviewers attending a QAR	905	943	1000
Number of reviewers trained	632	620	618
Number of training sessions taking place	33	35	40
% of complete review teams	87%	89%	88%
% of schools fulfilling their commitment	60%	56%	54%
Rating by host headteacher of the CPD for schools and their staff	4.6/5	4.6/5	4.7/5
Rating of the reviewers of their CPD	4.6/5	4.7/5	4.7/5
Rating of the reviewer training	4.6/5	4.6/5	4.6/5

Only 90 per cent of teams are fully staffed. This is because it is difficult to fill teams when staff drop out at late notice. There has also been an increase in the number of schools not fulfilling their quota of reviewers, due to demands on senior staff to remain within the school on the date they are assigned to a QAR and for personal reasons. Adding to this is the increasing proportion of small schools in Challenge Partners who generally find it more difficult to spare their few senior staff. All of these figures vary from hub to hub. Some hubs have a 100 per cent record for completed teams and a 90 per cent quota.

Ratings for the value of the professional development are high. This is corroborated by written comments.

Quote from a headteacher being reviewed:

> “More than anything it helps get your priorities in the right place” (Challenge Partners, p14, 2014b)

Quote from a reviewer:

> “You can use the QA Review as a learning process and it’s a way of trading the inside secrets” (Challenge Partners, p14, 2014b)

However, the average rating can mask the fact that on the rare occasions when the process is considered to have failed by the reviewers or the reviewed school the ratings drop significantly. Thus, a significant majority of staff involved rate the process higher than the average score.

8.19.3 External

In April 2015, Challenge Partners commissioned an independent evaluation of QAR (Mathews and Headon 2015). It used an extensive evidence base. Their key findings corroborated the high ratings of the process that Challenge Partners had and continues to collect from its schools. They reported that “the Challenge Partners’ QA review model is exceptional in its conception, rigour, quality and developmental power” (Mathews and Headon, p8, 2015).

They drew attention to the multiple gains to schools in the partnerships. These were:

- The checking of the reviewed school’s self-evaluation and the effectiveness of their school improvement strategies. 84 per cent of headteachers of reviewed schools indicated that reviews had been very useful to the professional development of their senior leaders.
- A critical analysis of ‘areas of excellence’ that schools are invited to nominate as part of their review which is then recorded in the subject directory, allowing other schools to access them. This rewards and encourages best practice.
- A focus upon areas of concern for each school and the provision of expert support locally and nationally to negate this.

- The exceptional professional development for the staff of the school being reviewed and the members of the review team.
- The discussion between senior leaders in the reviewer's home school of their experience and learning often leads to the adoption of improved practice and further communications or visits between not only the reviewed and reviewers' schools but also between the schools of members of the review team.

They concluded: "*The partnership approach to reviews, and the multiplicity of advantages that accrue from them, are distinguishing and possibly unique features of the Challenge Partners method.*" *(Mathews and Headon, p7, 2015)*

8.20 Conclusions

During the period of the case study between 2011 and 2018, the number of schools in CP has increased from 75 to 550. The exam performance of students has been consistently above national average both in attainment and progress and the gap between disadvantaged and average pupils has been smaller. Though the distribution of schools to OFSTED categories has varied, the rate of transformation has been consistently greater than the national average. Teaching schools have played a significant role in the development of Challenge Partners, especially at local level where their role in hub effectiveness is critical. Members of CP have played a leading role in the development of the school system.

It is apparent from the evidence presented above that the QAR builds upon the learning of a group of school leaders who have worked collaboratively over 15 years to build a theory of action focused upon how school-to-school work can maximise student achievement across a group of schools. The desire of the original group of schools that formed Challenge Partners to create a Peer Review that …

- Ensured participating schools challenge each other over performance and proposed strategies to help improve, and that they experience this process within their own and other schools.
- Gave reviewers time to reflect upon their own performance and learn from their peers.
- Identified areas of excellence which could be shared across the network.
- Identified schools requiring support or who could provide it after sharing the outcome of the peer review with their local schools and nationally.
- Played a major role in binding the collaborative network together through this process.

… has been achieved.

The description of the QAR is only for the period of the case study. In Challenge Partners' continual drive to improve the process, subsequent refinements have been made. These include the pilot of two modified QAR for high performing secondary schools, to be known as 'growing the top' and for Multiple Academy Trusts. Areas of excellence and the policies on QAR cancellation are also being updated, and information on the 'character and context' of the school and the reviewer training programme are being refined. An advanced reviewer role and training is being piloted, and a taxonomy is being introduced to ensure that an accurate picture of the needs and strengths of our schools and the information in review reports can be categorised.

It is within this continual developmental context that these conclusions are made.

High reliability in the QAR is achieved by learning from past practice and turning this learning into a set of non-negotiables and degrees of variance which are used to regulate the activity, standardising training for reviewers, selecting and training the lead reviewers, and monitoring the delivery.

They ensure that the QAR has validity as a peer review which makes a contribution to school improvement. It builds upon a theory of action which stems from best practice in school-to-school work and triangulating the school's judgments of where it is now and what it should do next against national criteria, along with the lead reviewer's own knowledge from acting as a trained Ofsted inspector and best practice.

Though the schools in Challenge Partners recognise that the QAR plays a key role in their improvement, it is only part of the improvement process. It assists the schools to define what needs to be done, where the support might come from and how they might support others to do the same. However, it does not guarantee the provision or implementation of that support. Here the ability of the leadership of the school to implement its plans and the capacity of the local hub and as required the central team to support them is critical.

Within the QAR process, the role the headteachers, the senior staff and the lead reviewers play in the action is critical. Their buy-in to the ethos of the partnership as an open knowledge managed network of schools determined to ensure every pupil maximises performance, their ability to build moral capital with their peers and their capacity to work and learn collaboratively is essential to the QAR being an integral part of a school's improvement.

Finally, Challenge Partners has always contained a group of schools from all phases which are among the higher performing in the country and have pioneered school-to-school work. Since its formation, it has provided a QAR to an increasing number of schools which perform collectively higher than the national average across a range of measures. The fact that all of these schools have been willing to pay for this service in a time of increasing budget pressures indicates the peer value they place upon it.

References

Berwick, G. T. (2011). The Approach: Engaging in Excellence, Volume One, London: George Berwick Associates.

Berwick, G. T. (2013). Moral Capital: Engaging in Excellence, Volume Two, London: George Berwick Associates.

Berwick, G. T. (2015). *Theory of action*. London: George Berwick Associates.

Berwick, G. T., & John, S. (2017). School networks, peer accountability and the brokerage of knowledge and expertise. In P. Earley & T. Greany (Eds.), *School leadership and education system reform*. London: Bloomsbury.

Challenge Partners. (2013). *Annual report April 2011 to August 2012*. London: Challenge Partners.

Challenge Partners. (2014a). *Reflection, challenge and development, peer reviews: The emerging picture*. London: Challenge Partners.

Challenge Partners. (2014b). *Annual report 2013*. London: Challenge Partners.

Challenge Partners. (2015a). *Making the most of the QA review*. London: Challenge Partners.

Challenge Partners. (2015b). *Annual report 2014*. London: Challenge Partners.

Challenge Partners. (2016). *Annual report 2015*. London: Challenge Partners.

Challenge Partners. (2017). *Annual report 2016*. London: Challenge Partners.

Challenge Partners. (2018). *Annual report 2017*. London: Challenge Partners.

Challenge Partners. (2019). *Annual report 2018*. London: Challenge Partners.

Department for Education. (2014a). Progress 8: How Progress 8 and attainment 8 measures are calculated. Accessed online on 3rd August 2020 at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/561021/Progress_8_and_Attainment_8_how_measures_are_calculated.pdf

Department for Education. (2014b). Teaching schools: a guide for potential applicants, Gov.UK. Accessed online on 3rd August 2020 at: https://www.gov.uk/guidance/teaching-schools-a-guide-for-potential-applicants

Department for Education. (2016). Understanding school and college performance measures, Gov.UK. Accessed online on 3rd August 2020 at: https://www.gov.uk/government/publications/understanding-school-andcollege-performance-measures/understanding-school-and-college-performance-measures

Department for Education. (2020). School and college performance measures, gov.UK. Accessed online on 3rd August 2020 at: https://www.gov.uk/government/collections/school-and-college-performance-measures

HM Government. (2010). *The importance of teaching: The schools white paper 2010, Cm7980*. Norwich: The Stationary Office.

Mathews, P., & Headon, M. (Eds.). (2015). *Multiple gains – An independent evaluation of challenge partners' peer reviews of schools*. London: IOE.

OFSTED. (2019). Education inspection framework (EIF), gov.Uk. Accessed online on 3rd August 2020 at: https://www.gov.uk/government/publications/education-inspection-framework

Professor Sir George Berwick CBE was one of the founders and Chief Executive Officer of Challenge Partners from 2011 to 2018. He was head teacher of an Ofsted rated outstanding school from 1993 to 2011, leader of the London Leadership Strategy strand of London Challenge, 2003-2008 and was CEO of Teaching Schools from 2008 to 2011. He has also worked in Canada and Sweden. He is credited with playing a major role in the development of school-to-school work from which the concept of Teaching Schools emerged. He is current a visiting professor at the IOE/UCL and Chairman of Olevi International.

Chapter 9
Education Development Trust's Schools Partnership Programme: A Collaborative School Improvement Movement

Andrew Ettinger, John Cronin, and Maggie Farrar

Abstract This chapter explores the background, principles, operations and evidence for and from the Education Development Trust's Schools Partnership Programme of peer review. This programme has been in operation for five years and has engaged with over 1,400 schools making it the largest peer review programme in England. The authors explore how peer review has emerged from, and is a central plank of, the English de-centralisation agenda and a bottom up desire for greater collaboration and peer networking to drive greater school improvement. The principles which govern the programme stem from a strong research base and are then connected directly to school leadership practice and reality. This then leads to a three-stage cyclical theory of change, which underpins the entire programme model and journey, leading to embedded peer review practice and genuine collaboration which in turn leads to greater teacher and school efficacy, and so better pupil outcomes. This sustainable maturity model means the programme has evolved with feedback and input from school leaders and the authors then look ahead to where next for peer review in England and internationally

9.1 Introduction

Education Development Trust (EdDevTrust) is an international charity whose mission is to transform lives by improving education around the world. Drawing on evidence-based practice and expertise they design and deliver effective and sustainable education solutions tailored to the local context. Collaborative school

A. Ettinger (✉) · J. Cronin · M. Farrar
Education Development Trust, Reading, UK
e-mail: aettinger@educationdevelopmenttrust.com; maggie@maggiefarrar.co.uk

© Springer Nature Switzerland AG 2020

D. Godfrey (ed.), *School Peer Review for Educational Improvement and Accountability*, Accountability and Educational Improvement,
https://doi.org/10.1007/978-3-030-48130-8_9

improvement and peer review is one of their core and growing areas of work in England and internationally.[1]

This chapter looks in detail at Education Development Trust's model of peer review, which has worked with over 1400 schools since 2014, making it one of largest collaborative school improvement programmes in the world. We will explore the rationale and principles that informed the development of the model and its theory of change for delivering continuously improving outcomes for young people. We will then set out the capacity, culture and systems that partnerships of schools engaged with the programme aim to embed at every level of their leadership to drive outcomes-focused improvement. Finally, we will reflect on the emerging impact to date and set out our aspirations for the next phase of development for the model.

This case study presents an approach to peer review that has been co-constructed with school leaders. The model has always been and will continue to be, developmental in nature, informed by the learning of the leaders that engage with the model every day. The reach and emerging impact of the model has led to it becoming the subject of a major national trial conducted by the Education Endowment Foundation and the Institute of Education.

9.2 An Education System Turning to Collaboration to Drive Improvement

The Schools Partnership Programme was developed in response to systemic changes in the school improvement landscape over the last decade. A de-centralisation agenda led to responsibility for school improvement shifting from local government to schools themselves. However, this greater autonomy went hand in hand with higher-stakes accountability. As a result, leaders increasingly looked to collaborate with other schools and peers – through Multi Academy Trusts, Teaching School Alliances, and more informal collaboratives – to provide the challenge, support and sharing of best practice needed to secure system-led improvement.

This emergent middle tier – when effective – can offer the school system a platform to enable targeted school improvement and the opportunity to support, share and embed improvement across schools (Mourshed et al. 2010). However, despite a long history of collaboration in England and a number of initiatives that enable school-to-school support, there is limited evidence about the improvement that partnership working actually delivers (Armstrong 2015).

---

[1] The authors of this chapter have all been closely involved in the co design, development and delivery of the model.

9.3 The Early Development of the Model

EdDevTrust's aim was to develop and sustain a self-improving model that strengthened and optimised the available capacity of the system to deliver effective and impactful partnership working.

It was this challenge that prompted EdDevTrust, together with headteachers from seven partnerships of schools, to find a new approach to collaborating for improvement. All the schools involved in this early phase had been previously involved in collaborative working and peer review, but they were all dissatisfied by the experience. Their experiences varied from peer review that felt too much like an inspection, or too loose and unfocused, or that met staff resistance when they felt 'done to', or had limited follow up and impact, or as one headteacher said at the time 'a lot of sound and fury but no change'.

9.3.1 A Strategic Question to Guide Change

Informed by the experience of these headteachers and our collective commitment to develop partnership working that can deliver improvement and positive outcomes for young people, we co-constructed a strategic guiding question we wanted to address together:

> How can we build on the strong commitment to collaboration in England to move from what could become a fragmented, diverse and potentially unequal system of schools to one which is collectively committed to continuously delivering improvement within their own and each other's schools?

We knew this long-term goal for the school system would rely on a new model of partnership working delivering a number of fundamental changes. Specifically those that create a shift from:

- Individual to connected autonomy;
- Partnership based collaborative projects to a collaborative improvement culture, characterised by strong relationships, high challenge and high support;
- Vertical (top down) to lateral (horizontal) trust based accountability achieved through building the agency for improvement at every level of the school and partnership;
- Individual and episodic, to collective and continuous review and improvement.

9.3.2 Peer Review as the Vehicle for Change

To help guide the initial development of the model we also invited Michael Fullan – a strategic friend to EdDevTrust – to draw on his work around effective collaboration. To deliver on the shifts we identified, Michael advised us to find what he referred to as a 'sticky vehicle'. This is something that would mobilise the greatest number of people across a partnership, embed a collaborative culture at partnership and school level, develop new forms of lateral trust based accountability and have a direct line of sight to better experiences and outcomes for teachers and pupils.

The sticky vehicle that we collectively identified was peer review. Our hypothesis was that when done well and as part of a cyclical improvement model including self review and school-to-school support, peer review could ultimately deliver continuous improvement. This would be through enabling schools to systematically review and address weaknesses and support the sharing of effective practice within and between schools.

The collective decision was taken to develop an approach to peer review that could effectively test this hypothesis.

9.3.3 The Guiding and Evolving Principles for the Model

The motivation for SPP was underpinned by a collective commitment and aspiration to secure school and system improvement through effective partnership working. Whilst the 'why' of SPP was always clear to us, it is only through constantly reflecting on what is working in practice with the hundreds of schools engaged with SPP that we are developing our learning of 'how' to make collaboration work.

The core principles underpinning the model and how these are embodied in its core components owes much to the work done by the initial group of seven partnerships and has remained the basis of the programme over the first 5 years of growth. In the first 3 years of development an advisory group was set up to provide expert scrutiny of both the core principles and the model. This group was made up of a core group of headteachers and others including Robert Hill (previous Government Adviser), Christine Gilbert (previous HMCI and Trustee of EdDevTrust) and David Weston (CEO, Teacher Development Trust). In addition, Michael Fullan, Pasi Sahlberg and Professor Vivianne Robinson offered support and challenge at key points in the programme's development. It was this group, drawing on their practice, expertise and research, that came up with the six core principles that guide the participants in their engagement in the programme.

These six principles have been refined over the years. In particular the annual 'review of reviews' challenge and support workshop that all partnerships engage in explores how the principles drive practice and what, if any, impact this has on practice and outcomes.

9.4 The Principles Underpinning the Model

The principles purpose is to ensure that host schools can expect their peer review to be conducted in a way that embodies specific behaviours and attitudes. They are central to the Theory of Change that drives the commitment to impact as a result of peer review (see Fig. 9.2) and are drawn from a school improvement research base.

The six principles are outlined in the following Table 9.1.

Table 9.1 Key principles for SPP

The principles	How these are exemplified in the model
Improve don't prove	A post review 'check in' which is probing and non-judgmental in style follows 3 months after the peer review to explore improvement and impact
Stay curious and open to enquiry	The practice of enquiry and curiosity, demonstrated through the practice of highly effective questioning throughout the review
Stay open to learning	Both the host school and the review team are equally open and willing to be changed by the experience, through detecting and challenging own and others assumptions, suspending judgment and giving and receiving honest and open feedback
Trust the evidence	The reviewers engage in evidence collection activities designed to respond to the area of enquiry set by the host school
Strengthen agency	Full staff engagement throughout the process agreeing the area of enquiry for the review and generating actions in the follow up improvement workshop
Practice reciprocity	A commitment to reciprocity demonstrated through the schools reviewing each other, where credibility comes from the skills of the peer review team, not the current performance of the schools they lead or work in

Each of these principles require the reviewers and the host school to adopt a set of behaviours and to strengthen particular attitudes regarding the purpose and practice of peer review. A commitment to an evidence-based approach has been an important element of the model.

Years of external review and monitoring of schools have tended to create a culture of 'putting on a good show', compelled by a need to 'prove' how good we are. This is understandable when staff are working hard and there is a need to demonstrate this when being reviewed or inspected. However the opportunity to learn can be limited in such an approach hence the principle of *'improve don't prove'*. This principle draws on the work of Cooperrider, D. L., Srivastva, S. (1987) on Appreciative Enquiry and the development of coaching cultures in organisations. This firmly positions peer review in a positive, strengths based framework, which is rigorous but supportive and more likely to lead to open and honest conversation and a commitment to make change.

This interest in peer review that is '*open to learning'* led us to spend time with Professor Vivianne Robinson exploring her work on 'open to learning conversations' (2018). Her insight that conversations about improvement are difficult because they have the potential to threaten relationships by generating discomfort and defensiveness rang true for many of our peer reviewers. This led us to strengthen the importance of peer review as an evidence collection process where assumptions are open to challenge and judgment suspended. It also helped us to shape the feedback session at the end of the review day as a dialogue about 'the findings' from the peer review between the host school and the reviewers in a way that strengthened the relationships, whilst acknowledging discomfort and feelings of defensiveness as they arise.

This means that the evidence base for the review, the evidence collection activities during the review and the way in which that evidence is analysed and narrated by the reviewers in the feedback conversations, is a critical part of the process hence the principle *'trust the evidence'*. Done well these collaborative and data rich conversations enable the reviewers to stay in enquiry mode and avoid advice or quick fix solutions. As well as drawing on 'big data' such as standardised tests, school inspections and surveys, Pasi Sahlberg through his work on 'good data' (2017) has worked with us to encourage the peer review process to draw on a range of evidence including "small data" using teachers' and students' "observations, assessments and reflections" of the teaching and learning processes in classrooms.

These data rich and evidence-based conversations allow all staff to engage in the review process in a way that feels authentic to them. Recent evidence from John Hattie (2016) suggesting that embedding collective efficacy across your school can be the biggest lever for delivering impact on pupils led us to position the principle of *'strengthen agency'* as even more central to the model than previously. This positioning was further influenced by the work of Donohoo (2016) suggesting that collective teacher efficacy can be developed through the practice of collaborative inquiry; questioning assumptions, challenging the status quo and provoking the difficult conversations needed to develop shared goals and effect change – all in fact reflected in the core practices of peer review within the SPP model.

As we developed the model over the last few years we have shared our approach with other systems practicing peer review in the UK and internationally. We have found that a review process can sometimes be called 'peer review' because practicing school leaders (i.e. peers) form part of the review team. However, in our view, peer review is termed such, not only because your peers are involved but because it is a reciprocal process, hence the principle of *'practice reciprocity'*. This is essential to develop the lateral trust based accountability that peer review can cultivate. Richard Elmore in his work on reciprocal accountability (2002) explored the impact of a mindset that believes that 'for every increment of performance I ask from you, I have an equal responsibility to provide you with the capacity to meet that expectation. Likewise, for every investment you make in my skill and knowledge, I have a reciprocal responsibility to demonstrate some new increment in performance.' This is Fullan's 'moral purpose in action' (2001) that has attracted so many of the partnerships we work with to peer review.

9.5 An Emergent Collaborative School Improvement Model

From this initial work, we developed a cyclical improvement model – with peer review at its heart – that could provide an approach for our partner schools to test together. Importantly it aimed to embed peer review into a collegiate and ongoing process of support and improvement that could produce the data and evidence on which focused and high impact school-to- school support could be provided.

It emphasises the fact that what happens before the peer review and what happens after it, is as important, if not more important than the review itself. Figure 9.1 demonstrates how each of the three phases of the programme support each other as

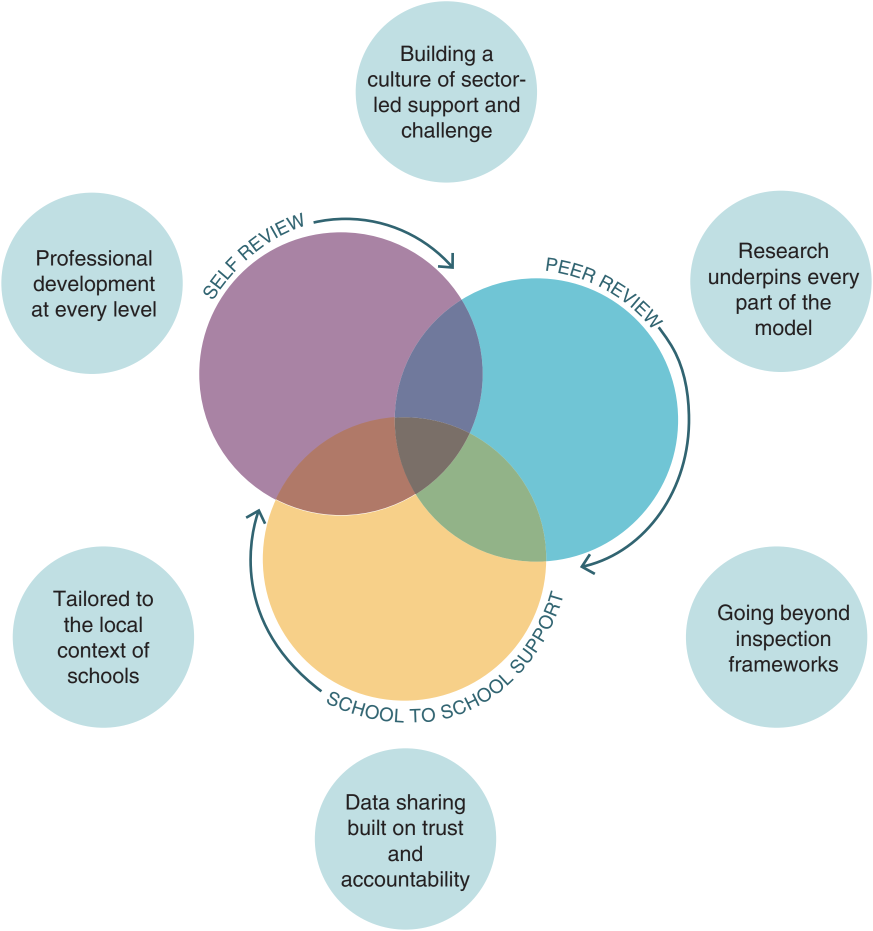

Building a culture of sector-led support and challenge
Research underpins every part of the model
Going beyond inspection frameworks
Data sharing built on trust and accountability
Tailored to the local context of schools
Professional development at every level
SELF REVIEW
PEER REVIEW
SCHOOL TO SCHOOL SUPPORT

Fig. 9.1 Cyclical improvement model

part of a continuous cycle of improvement, and the strategic approaches which ensure the model is embedded along the way.

The first seven schools developed and tested the early model in 2014, and since then new partnerships of schools have joined year on year. One of the core aims of SPP is to strengthen partnership, school, leader and teacher agency, and as partnerships join and make improvements to the model, these are then shared across the national network and become a core part of it. The programme is built on a theory of change intended to ensure that positive change happens and can be seen at scale.

9.6 Unpacking the Schools Partnership Programme; Our Theory of Change

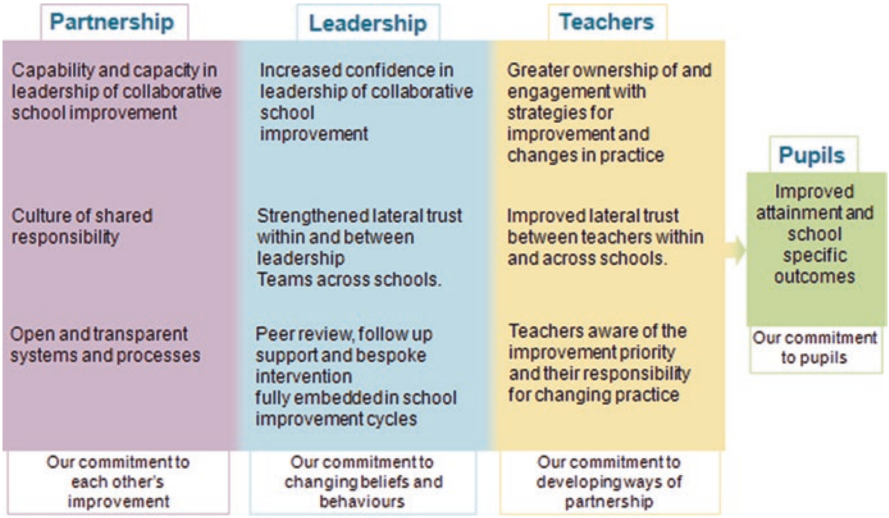
The theory of change underpins the cyclical improvement model and is intended to articulate how and why our model delivers change for partnerships of schools.

The top layer of the theoretical model in Fig. 9.2 demonstrates how the programme is built on enabling change at three levels – each level providing motivation and modelling of practice and behaviours to the others. The phasing of the programme builds out from the over-arching partnership through to the leadership in each school and then to the individual middle-leaders and teachers. Partnerships of schools keep their reviews under review, assessing progress and following up on school-to-school support and impact. Senior leaders lead school-to-school peer reviews while teachers and the wider school communities are involved in reviews as well as being trained to engage in middle leader and teacher reviews of practice and impact.

On a partnership level, defined in the first column, we expect to see impetus for change in the capacity of leading school improvement, creating an embedded culture of mutual trust and shared responsibility for outcomes, and developing open, transparent ways of working across the partnership and within schools.

Similarly, change is expected at three levels for schools and teachers as the programme moves through its three phases and the maturity of the partnership leads to a deeper understanding of, and more effective practice in, peer review for all – leading to genuine and deep collaboration towards shared goals.

The delivery of this change is supported by a training and development programme. This is intended to build the technical skills of peer review and school



Partnership	Leadership	Teachers	Pupils
Capability and capacity in leadership of collaborative school improvement	Increased confidence in leadership of collaborative school improvement	Greater ownership of and engagement with strategies for improvement and changes in practice	
Culture of shared responsibility	Strengthened lateral trust within and between leadership Teams across schools.	Improved lateral trust between teachers within and across schools.	Improved attainment and school specific outcomes
Open and transparent systems and processes	Peer review, follow up support and bespoke intervention fully embedded in school improvement cycles	Teachers aware of the improvement priority and their responsibility for changing practice	Our commitment to pupils
Our commitment to each other's improvement	Our commitment to changing beliefs and behaviours	Our commitment to developing ways of partnership	

Fig. 9.2 The SPP theory of change

improvement so all members of a partnership and local system can engage in regular scrutiny of each other's practice, gather and analyse evidence, give and receive feedback, and both provide and accept focused improvement support. On a partnership level the programme intends to develop a partnership culture characterised by a commitment by everyone in the cluster and local system to shared responsibility for improvement both within and between schools. This is achieved in a climate of openness, trust and honesty, and demonstrates a willingness to hold each other to account for agreed outcomes. Mature clusters agree one or two cluster priorities as a result of peer review, set agreed outcomes and hold each other to account for them.

Below the theory of change are four commitments that partnerships of schools involved in the programme make to each other. These commitments are expected to drive the agenda of partnership meetings and form the basis of the progress and impact telephone calls and workshops Ed Dev Trust conduct with the partnership. These commitments are to:

1. each others improvement through being willing to provide support following a peer review;
2. changing beliefs and behaviours regarding the importance of collegiate approaches to school improvement, teacher development and high quality learning experiences for children and young people;
3. developing effective school and partnership based systems and structures that enable shared improvement planning and the effective training, development and deployment of staff;
4. better outcomes and life chances for our own and each others children and young people.

The desired changes illustrated in the theory of change are aligned with the core programme for schools. This is intended to develop an explicit awareness and engagement with the changes schools and leaders are trying to effect through the model.

The next section will explore in detail the key stages and principles of the SPP model that have been refined over the last 5 years.

9.6.1 Stage 1: School Self-Review and the Peer Review Framework

In order for the host school to know what it wants the peer review to focus on, the process starts with a school self-review using the peer review framework. In order to ensure that peer review is not an 'add on' process, schools who have engaged in the model successfully have embedded it into their current management systems, and will therefore position the timing of the peer review in the school improvement planning cycle. Self-review is most effective when all the staff engage in the process and have a hand in agreeing the focus for the peer review, this accords with one of our core principles of ‘strengthening agency’ (see Table 9.1).

This is the first time the peer review framework will be introduced into the process. This was developed by the first seven partnerships and then refined over the following 3 years. It takes account of the core elements in the OfSTED framework but schools were very keen to have a framework for improvement that they owned, which was not an inspection framework.

The initial framework had a series of statements and ‘look fors’ to guide the peer review team but after it was reviewed by a senior HMI, who criticised it as being ‘statement and not enquiry based’ and ‘limiting not enhancing the agency of school leaders’, it was reworked to be based on a series of questions rather than statements. It is organised under four themes. Each theme has three dimensions. Each dimension then has a series of enquiry questions beginning ‘ to what extent ……’ to enable the host school to come to an agreement on the most useful enquiry question/s for the reviewers (Fig. 9.3).

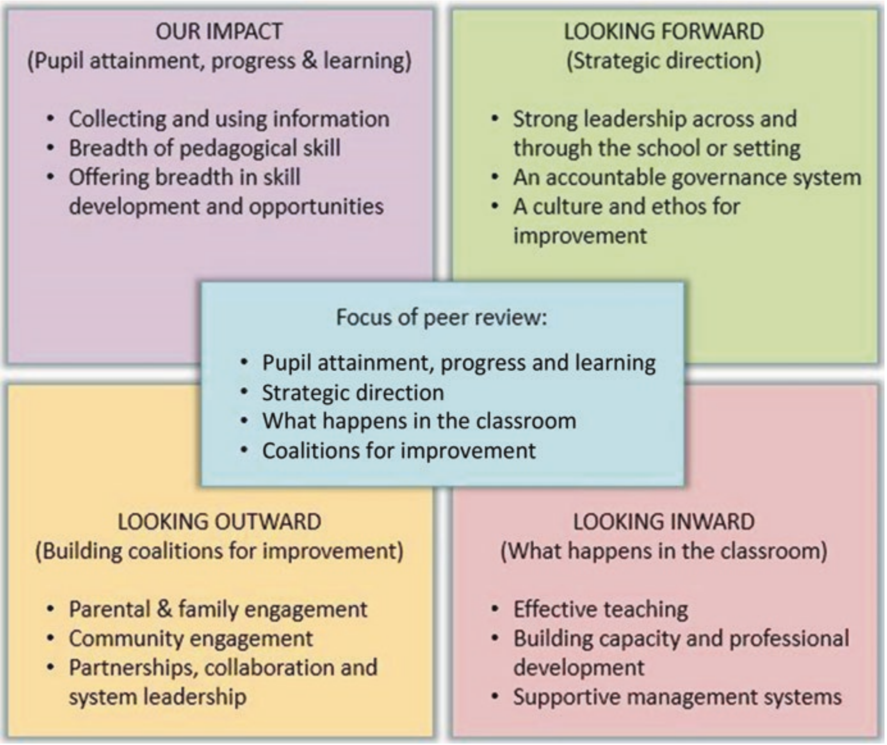

OUR IMPACT
(Pupil attainment, progress & learning)
• Collecting and using information
• Breadth of pedagogical skill
• Offering breadth in skill development and opportunities
LOOKING FORWARD
(Strategic direction)
• Strong leadership across and through the school or setting
• An accountable governance system
• A culture and ethos for improvement
Focus of peer review:
• Pupil attainment, progress and learning
• Strategic direction
• What happens in the classroom
• Coalitions for improvement
LOOKING OUTWARD
(Building coalitions for improvement)
• Parental & family engagement
• Community engagement
• Partnerships, collaboration and system leadership
LOOKING INWARD
(What happens in the classroom)
• Effective teaching
• Building capacity and professional development
• Supportive management systems

Fig. 9.3 The SPP peer review framework

9.6.2 Stage 2: Peer Review

An initial conversation between the lead reviewer and the host school refines the focus, which is agreed by the host school and is the focus of enquiry and evidence collection by the review team. The role of the team is to make sure that this focus is based on evidence of what needs to improve and what outcome is going to be of most benefit to the host school.

The reviewers and the host school look at the outcome of the self review and consider the relevant enquiry questions. An initial conversation between the reviewers and the headteacher explore what 'great' would look like in the relevant area in the context of the school, and what wider evidence the headteacher draws on to give that view. These conversations embody the principles of enquiry and agency, enabling the host school to set the benchmark for improvement and for the peer reviewers to then collect evidence to assess the 'as is' situation against the benchmark of future improvement set by the school.

The overall aims of the review team are to:

- assess the validity of the self-review process through examining the school's own assessment of its strengths and areas for improvement;
- identify areas of inspirational, excellent and effective practice that will be of benefit to the school and the partnership;
- strengthen the school's own capacity for self-review through its engagement with the review team;
- identify areas for development that will be explored in the follow up improvement workshop and met either through the school's own approach to improvement and/or through follow-up brokered school-to-school or teacher to teacher support.

The peer reviewers then spend a day in the host school to enquire into the areas agreed by the school. Data collection takes place through a number of approaches– according to the focus of the review – such as learning walks, interviews with individual or groups of staff or students, closer analysis of school data, etc. Our expectation is that peer reviewers work as a team or trio in order to triangulate evidence and jointly analyse their findings. They are not there to pass judgment but to seek evidence and agree findings that are then shared with the school. As peer review is also a professional development opportunity, we also encourage a middle or senior leader in the school to take the opportunity to shadow the review team and give them feedback at the end of the process.

Immediate feedback at the end of the review day is important, as is coming to a shared agreement on what needs to improve. For this reason, the peer review ends with a feedback conversation. This conversation will describe the evidence gathered with any main findings, and as an appreciate enquiry approach will always include reference to elements to celebrate and will also identify an improvement priority for the follow up workshop.

9.6.3 Stage 3: Follow Up Workshop and School-to-School Support

In order to ensure that the review findings are then turned into action, the SPP model includes a post-review improvement workshop that takes place no more than 2 weeks after the review. The lead reviewer attends (other review team members may also do so) and the headteacher of the reviewed school decides who should attend the workshop, depending on the focus and outcome of the review.

For most schools this takes place in a staff meeting and, using a range of facilitation tools, is designed to get to the root cause of issues, agree actions and broker any necessary support from other schools in the cluster.

Middle leaders and aspiring senior leaders across the cluster are trained as Improvement Champions to facilitate the improvement workshops in schools other than their own. They are also trained in initial research skills to identify research and evidence to inform the workshops and strengthen collaborative school improvement by identifying priorities for improvement across the partnership.

The most important aspect of their role is supporting the school staff to take ownership of the outcomes of peer review in the improvement workshop. Through their training and the facilitation techniques and tools they are provided with, Improvement Champions ensure shared solutions are produced in the workshop and that actions are clearly delegated across the staff body.

School-to-school support follows the review process and arises from specific areas of need identified through the action plan. The partnership lead and/or the Improvement Champion can play a key role in brokering appropriate support by drawing together a team from across the partnership in order to meet the support needs of the school. The team should comprise people with the necessary skills to provide high quality school-to-school support. As agreed at the improvement workshop, three levels of school-to-school support may be required:

1. Within the reviewed school – with colleagues working together to support and coach;
2. Across the partnership – with colleagues from another partnership school supporting a colleague in the reviewed school;
3. With colleagues from outside the partnership, to provide bespoke support.

In all cases, schools in the partnership need to agree their terms of engagement (or contract)' and protocols for all staff involved in the support together with practical issues, such as who will be involved and how often.

As well as keeping peer review under review at the partnership meetings the model also includes a 3 month check-in following the improvement workshop conducted by the lead reviewer and one of the Improvement Champions who facilitated the workshop.

This follows a common process of checking on actions and follow up, looking at evidence of quality and then evidence of impact.

9.7 Long Term Sustainability

For peer review to be embedded into the ways of working of schools and partnerships the partnerships need to keep it under review so they can continually refine and further embed the process.

The SPP model recommends that all partnerships undertake an annual review of reviews workshop to both assess impact and agree improvements for the following year. Those that have done so are starting to agree shared enquiry areas of focus for the peer review and to strengthen their follow up school to school support processes, as well as agree how new staff will be inducted into the process with appropriate shadowing opportunities and training.

Partnerships that have been with the programme for 3 years or more are also developing their own internal Quality Assurance (QA) teams. Experienced peer reviewers work in pairs and step out of the process for a year at a time to sample and QA the review process. This process was originally co-constructed with a group of leaders in the North East of England. They agreed that the purpose of QA was to:

- assess how effectively and consistently peer review is operating across the partnership;
- increase the reliability and consistency of peer review which in turn will ensure impact;
- identify areas of strength within a partnership's practice of peer review that the partnership can share;
- identify areas of development and improvement in the practice of peer review that can then be acted on;
- identify the strong peer reviewers so that the pairing of reviewers in future reviews can be done strategically to support their development as reviewers.

The QA framework is built around the six core principals and the processes underpinning the model.

9.8 So What? Peer Review, Partnership Maturity and School Improvement

Following 5 years of development and engagement with over 1400 schools, the impact of the model is beginning to be identified at both micro level (what are individual partnerships and school leaders are saying) and macro level (what impact is it having on school outcomes).

On the micro level, school leaders and partnership leads have given feedback on the purpose of peer review in their locality and the difference it is beginning to make.

The Kyra Teaching School Alliance in Lincolnshire describe their experience as a 'journey', and one that required them to be braver in examining the impact of what they do. Their strategy was built around a need to legitimise the necessary but tough conversations about the quality of relationships and about outcomes. Marie Claire Bretherton, who leads the Alliance describes this as a collective commitment to use peer review as an opportunity to ask questions focused on the child and their experience:

> Let's put the child at the centre of what we are doing and ask questions about how well that pupil is learning, what's the pupils' experience of the schools and the curriculum in that particular setting, and use that as a kind of leveller to answer questions about pedagogy, about curriculum, about leadership, about assessment. For us, it's raised expectations, it's normalised school-to-school support, and it's normalised the use of expertise across the system.

Liz Robinson at Surrey Square Teaching School shares her own insights into this shift in the quality and type of conversations that take place through peer review saying *'Peer review gives us the legitimacy to have those conversations we know we have needed to have for some time but didn't quite find the right time, space and words to have them'.*

Similarly Helen Rowland of Focus Trust highlights the quality and type of questions that are legitimised by peer review as a key and valuable part of the process, leading to the discovery of new and potentially challenging information:

> Questions were asked that had not been considered (or that had been avoided) before in a very challenging way. Although some of the information revealed was not what we wanted to hear it was shared in a constructive way and we felt positive about ways to move forward.

The Centurion Partnership in Lincolnshire has engaged in peer review as a process to strengthen the maturity of the partnership. Previously the schools collaborated around events such as sports competitions and activity for children with some shared professional development for teachers. Peer review has shifted the relationships and the purpose of the partnership to focus on collective improvement and impact and this in turn has changed the nature of their partnership meetings, as described by Ian Tyas:

> Our meetings have become much more about leadership and pedagogy, and importantly about how we can support each other in our own school's journey.

This shift to a focus on collective knowledge of each others schools in a partnership and a commitment to work together for improvement is echoed by Richard Potter, Chair of the Colchester Schools Consortium, with a particular emphasis on knowing about and capitalising on effective practice:

> Like most in the teaching profession, there is a reluctance to promote 'good practice' due to modesty and sometimes capacity to share that practice, but the peer review process highlights pockets of good practice or good practitioners to the steering group and so different schools' strengths are known and mapped over the cluster. If challenged, the steering group and to a large extent the cluster heads as a whole, could speak knowledgably about any school within the cluster based on the feedback that has been shared out of peer review.
>
> Each school links their school improvement plans to the cluster improvement plan enabling common partnership wide aims to be agreed. These aims then allow the partnership to plan actions based on peer review outcomes and school needs. Peer review then becomes an integral monitoring as well as support tool for the schools and steering group alike.

The Primary 6 Partnership in Nottingham cites the development of a shared and common language through which to talk about improvement and hold themselves to account for taking action as an important outcome from peer review. As the peer review process matures, more staff engage in the practice making school improvement everyone's responsibility

> We have developed a language of challenge, helping to hold ourselves collectively to account for subsequent actions. We have found new ways of engaging with staff at every level so there are more people asking, "'how can this part of school life improve?"'

These experiences of peer review in individual partnerships, and the emerging outcomes we are witnessing supports the message that we share with those expressing an interest in peer review. It isn't a project. It isn't an activity. It is a process that can strengthen the maturity of a partnership, change ways of working to focus on impact and outcomes, and strengthen the culture to enable shared accountability for improvement to flourish.

On the macro – system level, our analysis of Ofsted performance for schools engaged in SPP for at least 1 year (Farrar and Cronin 2017) we found that schools are improving against their baseline Ofsted inspection grades upon joining the programme. The data shows that schools engaged in the programme are significantly more likely to improve by one or more grades in inspection than the national average – 67% vs 61.8% (based on 500 schools). Although it is acknowledged that this is a blunt analysis and other factors may be in play, it provides a basis for further investigation. Additionally then, a summary of our annual review of reviews workshops with partnerships reports that schools have greater confidence to talk about their practice to 'visitors', they have improved confidence in analysing evidence and the increased engagement of the whole staff in the improvement action planning following peer review, have played a part in their improvement.

In a largely rural County in the East of England a stubborn 'flat-lining' on government spelling tests was investigated through peer review leading to a rise in standardised spelling scores from 62.2% to 90.4% by the end of the year. In another school in the same area where there the peer review focused on variability in outcomes in mathematics, these have improved over the 2-year period since the first review from below to above national average.

At a Trust in the North of England an improvement priority on improving children's reasoning in mathematics gave staff the opportunity to work together to develop a whole school approach to mathematical reasoning in children from 3 to 11 years of age. This has proved so promising that it has now been shared with 15 other schools and forms part of the professional development programme offered to all newly qualified teachers joining the Trust. In addition the practice of peer review and the monitoring of the improvement priorities has led to a tightening of the partnerships' priorities to 'The Focus Five'.

In another part of the country, a school in Birmingham reported that as a result of peer review children are receiving a more 'personalised' maths education. Effective questioning is now identified as a strength in the school and was previously noted as an area to be addressed in the improvement workshop, following the peer review.

These improvements are in part coming about as a result of school leaders inviting peers with a ‘fresh pair of eyes’ into their school, to gather evidence and respond to an enquiry question that they and their staff had posed. What actions the school took as a result of the peer review were then agreed in the highly active and facilitated improvement workshop where all staff had a hand in agreeing the solution and made a collective commitment to changing their practice as a result. In many cases this required the school leaders to ‘step back’ and let their staff lead the improvement, and in some cases this is what has made significant and lasting change possible.

9.9 Five Years On; the Growth of the Model

Since the pilot of the model in 2014, over 1400 schools across 120 clusters have engaged with the SPP 3-year maturity model.

Schools who join become part of a national and collegiate network of partnerships of schools engaged in peer review with opportunities to influence the development of the model and learn from each other’s practice.

Since the early development of the model further thinking on effective collaboration has given impetus to its growth. Michael Fullan and Steve Munby, in writing about the emergence of a new ‘middle tier’ in education, (2016) note that one of the critical success factors for system-wide school collaboration is: ‘a commitment to, and capacity for, effective peer review which forms the engine that drives improvement’.

Ownership is one of the core principles of the model and Education Development Trust is committed to giving schools the ability to lead the model themselves. A national network of ‘peer review hubs’ led by experienced partnerships has been established. Firstly, to provide national advocacy and influence, as well as providing strategic direction on the development of the model; and secondly to support local growth, training and new clusters across a region.

By involving schools so closely in the ongoing development and direction of the programme, we enable them to assume true ownership of the model and own its development year on year. The Theory of Change underpinning the model was developed in the third year of growth, in collaboration with school leaders. The aim was to develop a model that exemplified what we were trying to achieve through peer review and how the model enabled that to happen over a 3-year period.

9.10 Looking Ahead

Peer review is still in its infancy. Its power and potential is still emergent. The SPP model continues to be tested and is adaptive, learning from the partnerships to continually improve it. Some schools remain nervous of the exposure that review by your peers can bring particularly in a market-led competitive education system. Some schools – although very few – feel they have nothing to learn from other schools. One of the key pieces of learning from the model to date is that a peer review model cannot be forced on a group of schools. An imposed model runs the risk of schools going through the motions but gaining little benefit.

The next stage in the SPP model is the development of partnership-to-partnership peer review. This would provide a networked approach to improvement at a regional and national level. In order for partnerships to deliver systemic impact, leaders must be committed to addressing the unevenness in outcomes in their local and regional systems.

There is a powerful 'drumbeat' underpinning effective peer review. It is not episodic, something that's done and ticked off on an annual calendar. It is fundamentally a 'way of being' for groups of schools that will change the way they work together.

As the school system becomes more mature it will gradually become more effective at the core functions of review, intervention and improvement. As schools become more autonomous and more accountable, peer review creates a climate and a culture where connected autonomy and trust-based accountability can grow. Schools engaged in SPP are now increasingly extending the practice to middle leaders and teachers within and between schools, so that the cycle of collaborative improvement is embedded at every level.

However, peer review could just as easily become something else. It could become hijacked by government and mandated. It could become little more than cosy chats in each other's schools. It could become a model that isn't reciprocal but 'done to' schools. It could become one where the conversations that need to happen don't happen because it's hard. It could become one where we review but don't improve each other's schools.

As more schools engage in the process and we facilitate the annual partnership 'review of reviews' we are learning about the fragility of the process in some areas and the risks to sustainability and further impact that need to be addressed. These include challenges such as competing priorities which distract from the relentless focus on agreed peer review priorities; that isolation can affect clusters of schools, and that peer review needs to be fully integrated with national and local accountability systems.

Our view is that, done robustly and rigorously, peer review forms the backbone of trust based and lateral accountability that can co-exist with top down and regulatory forms of accountability.

We believe – and this book is a timely contribution to that debate – that there is a growing number of leaders prepared to invest in reciprocal peer review because they believe it's the right thing to do and it gets results. They want to reclaim what it means to be a great school, and to have the necessary conversations with each other about what needs to improve.

Above all, the leaders we have had the privilege to work with, know that peer review makes a reality of collective moral purpose. They model this through being willing to hold themselves and each other to account for improvement, through being ready to make their best practice available to each other across the cluster and through their willingness to tackle issues of collective importance that will ensure that the greatest number of children and young people benefit.

References

Armstrong, P. (2015). *Effective school partnerships and collaboration for school improvement: A review of the evidence*. London: DfE.

Donohoo, J. (2016). *Collective efficacy: How Educators' beliefs impact student learning*. Thousand Oaks, California: Corwin.

Farrar, M., & Cronin, J. (2017). *The power and potential of peer review.* EdDevTrust. https://www.educationdevelopmenttrust.com/en-GB/our-research/our-insights-library/2017/power-and-potential-of-peer-review. Accessed Dec 2018.

Fullan, M. (2001). *Leading in a culture of change* (Rev ed.). San Francisco: Jossey-Bass.

Fullan, M., & Munby, S. (2016). *Inside-out and downside-up how leading from the middle has the power to transform education systems*. Motion-Leadership & EdDevTrust, online resource. https://michaelfullan.ca/wp-content/uploads/2016/06/EdDevTrust-Global-Dialogue-FINAL.pdf, https://michaelfullan.ca/wp-content/uploads/2016/06/EdDevTrust-Global-Dialogue-FINAL.pdf. Accessed Dec 2018.

Hattie, J. (2016). *Third visible learning annual conference: Mindframes and maximizers*. Washington, DC, July 11, 2016.

Mourshed, M., Chijioke, C., & Barber, M. (2010). *How the world's most improved school systems keep getting better*. London: McKinsey & Company.

Robinson, V. (2018). *Open-to-learning conversations: Background paper introduction to open-to-learning conversations*. New Zealand: University of Auckland.

Sahlberg, P (2017). *Small data for big change: What does it mean for teachers?* Pasi Sahlberg Copyright. https://pasisahlberg.com/small-data-for-big-change-what-does-it-mean-for-teachers/. Accessed Dec 2018.

Andrew Ettinger is Head of Collaborative School Improvement at the Education Development Trust – there he leads the national School Partnership Programme, as well as over-seeing the organisations' wider collaborative school improvement strategy and portfolio. Prior to this, Andrew was Director of Leadership Development at Ambition School Leadership, leading the design and delivery of the Teaching Leaders and Future Leaders programmes and Ambition's NPQ qualifications. Before joining Ambition, he was a policy advisor at the Department for Education with responsibility for enacting the Raising the Participation Age legislation in England and a professional officer at the Teacher Training Agency leading on school and ITT provider partnerships.

John Cronin oversaw the initial development and national roll-out of the Schools Partnership Programme. Starting with a small pilot in early 2014, John helped grow SPP to work with over 1,300 schools across England, including more than 450 schools through the Education Endowment Foundation funded evaluation of the programme. John previously managed educational projects around school improvement and evaluation both nationally and internationally.

Maggie Farrar works nationally and internationally on partnership and networked based approaches to school improvement and leadership development. She was previously a senior Director at the National College for School Leadership and is currently working as an Education Consultant. She is an Associate for Education Development Trust.

Part VI

Participatory Evaluation Approaches to Peer Review

Chapter 10
Empowering Principals in Peer Review: The Value of an Empowerment Evaluation Approach for Educational Improvement

Kerrie Ikin

Abstract This chapter describes through the eyes of the participating principals the findings of a 3-year school peer-review research project conducted within the New South Wales government education system. The research compares an empowerment evaluation approach with a participatory and collaborative one. While both these stakeholder engagement approaches to evaluation changed the traditional practice for school accountability and improvement reviews and influenced the principals to change their perceptions of and attitudes towards reviews and professional accountability, including the way review processes and results were used in their schools, the research indicates that an empowerment approach is superior for building principals' evaluation capacity and improving organizational learning. The findings also show that there are a number of factors that are necessary for change, and even more importantly that there are underlying values, such as trust, openness, and transparency, that act as catalysts for change.

Evaluating the performance of schools is not new—either in New South Wales, or more widely across Australia, or in other countries. Mechanisms for the evaluation of school performance and their balance between accountability and improvement have long been the subject of research and debate. In an era of global educational reform, peer-review processes have emerged as an important way to address school improvement and professional accountability. Possibly a little ahead of its time, in 2006 an innovative peer-review program was developed, piloted, and subsequently implemented in one Region of the New South Wales government education system. Based on collaborative, participatory, and empowerment evaluation practices and theory (Fetterman et al. 2018b, d), the aim was to engage principals as evaluators of their own and their colleagues' schools in ways that boosted evaluation capacity across all their institutions and led to school improvement.

K. Ikin (✉)
University of New England, Armidale, Australia

© Springer Nature Switzerland AG 2020

D. Godfrey (ed.), *School Peer Review for Educational Improvement and Accountability*, Accountability and Educational Improvement,
https://doi.org/10.1007/978-3-030-48130-8_10

This chapter presents the findings from a 3-year research project that was undertaken during the development and pilot phases of this program. It provides an overview of the program and how it was developed, how various components of empowerment, collaborative, and participatory evaluation approaches were applied to the program, and how the research was undertaken. Most importantly it describes the process and its impact through the eyes of the 18 participating principals, the influences that participation in this peer-review program had on them, and the lessons learnt. It concludes by making a case for the use of empowerment evaluation practices using within-school peer review in all schools.

10.1 The Context in Which the Program Emerged

Reviews had long formed a part of the educational landscape in New South Wales. From as early as 1848 school reviews, in one form or another, were integral to the accountability function of the New South Wales government school education system, and elements of even the earliest reviews still remained in review processes at the beginning of the twenty-first century. Also forming an integral part of the educational landscape in New South Wales were the voices of the New South Wales Teachers Federation and the two groups representing principals. The principals' groups, while not trades unions, had nevertheless increasingly held sway in the accountability function in schools, particularly where the supervision and accountability of principals were concerned.

The Scott Report of 1990 (New South Wales Education Portfolio 1990), the subsequent Quality Assurance era, and the emergence of new school-accountability frameworks from the mid-1990s, coupled with growing action by the New South Wales Teachers Federation and advocacy by the two principals' groups began to change this emphasis. Despite this, the views of and leadership by external experts (most often external to the school but internal to the education department and more senior to the school principal) were still seen by the system as late as 2005 as a necessity for school review. Further, the regular review of all schools, although mooted by the education department, had not been agreed to industrially. Reviews took place only when a problem in a school came to the attention of the director responsible for that school; that is, reviews were by exception (New South Wales Department of Education and Training 2004).

It was not until the introduction of the Cyclical Review program in 2006 that principals were given responsibility for leading a school review for system accountability.

10.2 The Brief for the Cyclical Review Program

The brief was to strengthen and complement the system-wide 2005 school development and accountability framework. The program was to provide a mechanism that would allow each school to be judged approximately every 4 years against statements of exemplary practice that reflected school operations and performance at the highest levels. It was envisaged that these reviews would have a long-term impact, in contrast to the existing reviews-by-exception, which, anecdotally, tended to produce an initial flurry of activity that soon dissipated.

The program also had to focus on four core strategy areas: (1) developing a mechanism for school evaluation that would be a robust, useful, and influential evaluation of whole-school performance and governance for the school being reviewed; (2) providing a sustainable regional framework in terms of time, personnel, and cost for such reviews; (3) building the evaluation capacity of principals; (4) building a culture of trust with regard to the reviewing of schools. The long-term goals were that the Cyclical Review program should be recognized as the model for developmental and accountability evaluations in government schools[1] in the State, and that it should not only be accepted but eagerly sought by principals. (Both the New South Wales Teachers Federation and the two principals' groups had long been against any form of compulsory school review. The idea was to make the Cyclical Review program so beneficial that most principals would volunteer to be part of it.)

10.3 The Peer-Review Approach Adopted for the Cyclical Review Program

The brief for the program required a sea-change in the way school accountability and development was viewed by principals, staff, the system, and the community and therefore a sea-change in the way school reviews had operated until this time. Collaborative, participatory, and empowerment evaluation approaches—all stakeholder involvement approaches—have substantial underlying guiding principles, values, skills, practices, and methods in common that were appropriate for the task (Fetterman et al. 2018c, d). These include a participant focus, learning, community knowledge, trust, ownership, and flexibility. They also have distinct differences, particularly regarding the stakeholders' role: how the evaluation is controlled and by whom. In empowerment evaluation, the stakeholders are in control of the evaluation, with the evaluator acting as a critical friend. In participatory evaluation, the evaluator and the stakeholders jointly share control of the evaluation. In collaborative evaluation, the evaluator is in control, but ensures that there is continuous engagement between the evaluator and the stakeholders (Fetterman et al. 2018b). These differences can lead, amongst other things, to different levels of inclusion, capacity building, and organizational learning.

---

[1] The School Cyclical Review Framework was recognized by the Australasian Evaluation Society Inc. in its 2008 Award for Excellence in Evaluation, winning the Best Evaluation Policy and Systems Award.

The program comprised three elements. The first element of the program was the formation of an advisory group of nine principals (Group 1) and the Region's senior school accountability and improvement officer (hereafter called the 'critical friend'), who had expertise in program evaluation. This group worked together to develop every aspect of the Cyclical Review program: the exemplary practice statements on which schools would be judged, the principles that would guide the program, how the reviews would operate, the composition of the review teams (Table 10.1), the tools that would be used to collect the data, and how the data would be recorded, analysed, and reported. In other words, the principals developed the program themselves for themselves and their peers. The process and timeline for developing the Cyclical Review program described in this paragraph is shown in Fig. 10.1.

The second element of the program was the development of actual peer reviews. Group 1 principals, with advice and assistance from the critical friend, where required, developed and led just-in-time training for review team members. Each review team included a further principal from another school. These nine principals formed Group 2. The Group 1 principals both reviewed one of the other schools of their Group 1 peers and had their own school reviewed. That is, they acted as both a review team leader and a host school principal. The Group 2 principals only took

Table 10.1 Composition of review teams

To develop a sense of ownership of the process and build the evaluation capacity of many more school and regional staff and, in particular, the principals, each review team comprised:	
Group 1 Principal	a principal of the host school, considered part of the team—collaborating with the team-leader principal in the organization of the review, providing additional information or clarification throughout the review as required, and taking part in the discussions about the analysis of data
Group 1 Principal	a principal as the team leader
Group 2 Principal	a principal from another school in the region not necessarily closely connected to the host school
School Staff	School staff, including executive staff, from within and beyond the school (the number of team members depended on the size of the school being reviewed. The number ranged from 4 to 9; 4 for the smaller primary schools and up to 9 for the large secondary schools)
Regional Staff	a regional consultant, selected to match the perceived development needs of the school
Other Personnel	For example, parents, personnel from special interest groups—could be members of the review team, depending on the needs of the school
Critical Friend	The senior school accountability and improvement officer (experienced and trained in conducting school reviews) as an ex-officio team member, acting as the critical friend to the team

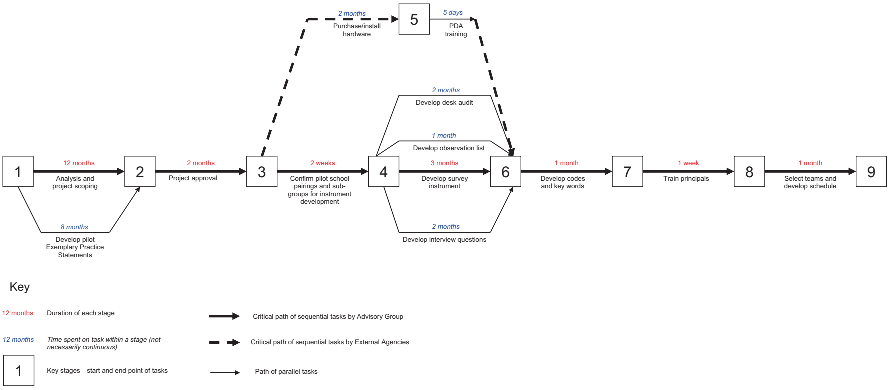

2 months
Purchase/install hardware
5
5 days
PDA training
2 months
Develop desk audit
1 month
Develop observation list
1
12 months
Analysis and project scoping
2
2 months
Project approval
3
2 weeks
Confirm pilot school pairings and sub-groups for instrument development
4
3 months
Develop survey instrument
6
1 month
Develop codes and key words
7
1 week
Train principals
8
1 month
Select teams and develop schedule
9
8 months
Develop pilot Exemplary Practice Statements
2 months
Develop interview questions
Key
12 months
Duration of each stage
12 months
Time spent on task within a stage (not necessarily continuous)
1
Key stages—start and end point of tasks
Critical path of sequential tasks by Advisory Group
Critical path of sequential tasks by External Agencies
Path of parallel tasks

Fig. 10.1 Process and timeline for developing the cyclical review program

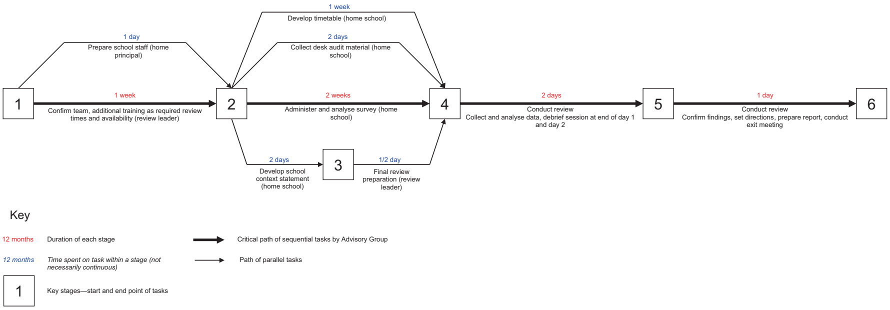

1 week
Develop timetable (home school)
1 day
Prepare school staff (home principal)
2 days
Collect desk audit material (home school)
1
1 week
Confirm team, additional training as required review times and availability (review leader)
2
2 weeks
Administer and analyse survey (home school)
4
2 days
Conduct review
Collect and analyse data, debrief session at end of day 1 and day 2
5
1 day
Conduct review
Confirm findings, set directions, prepare report, conduct exit meeting
6
2 days
Develop school context statement (home school)
3
1/2 day
Final review preparation (review leader)
Key
12 months
Duration of each stage
12 months
Time spent on task within a stage (not necessarily continuous)
1
Key stages—start and end point of tasks
Critical path of sequential tasks by Advisory Group
Path of parallel tasks

Fig. 10.2 Process and timeline for conducting a cyclical review

part as team members in one peer review. The process and timeline for conducting a Cyclical Review described in this paragraph is shown in Fig. 10.2.

The third element involved the Group 1 principals acting in their advisory group role to consider their experiences in and feedback from each review, so that continual adjustment and improvements could be made. They further acted as critical friends themselves to new review teams. This element also involved the establishment of a steering committee to oversee full implementation of the Cyclical Reviews. In this way the program itself became cyclical, providing mechanisms for a continuous cycle of improvement.

Thus, the approach adopted for Group 1 principals was empowerment. They developed and controlled the evaluation process, with a trained evaluator acting as a critical friend. The approach adopted for Group 2 principals was a mix of participation and collaboration. They were stakeholders, not only because they were members of review teams but also because it was envisaged that, in time, they would themselves lead reviews and have their own schools reviewed. As team members, they jointly shared control of the review itself once it got under way. This has similarities with a participatory approach. Group 2 principals did not, however, develop the process. At the same time there was nevertheless continuous engagement between them and the trained evaluator about the process and its effectiveness. This has similarities with a collaborative approach.

10.4 The Case-Study

Given the historical context of school reviews and the desire to develop a process that would eventually permeate the Region and possibly the State, the decision was made to undertake research to identify the impact on the participating principals and also provide continuous knowledge to the advisory group and steering committee so that the program could be modified and strengthened.

The research focused on three key questions.

1. What factors, prior experiences, and understandings contribute to the influence that involvement in Cyclical Reviews had on the participating principals?
2. How does participation in Cyclical Reviews influence participating principals?
3. To what extent are the outcomes of evaluation capacity building (Baizerman et al. 2002a, b) demonstrated by the principals who participated in the Cyclical Reviews?

A qualitative approach, using a case-study technique (Yin 2003) for the gathering of data from multiple sources, was seen as the most appropriate method to answer these questions. The case-study itself examined the two principals' groups: Group 1 ($n = 9$) and Group 2 ($n = 9$).

Both groups had similarities in their composition. Each group had both male and female principals from both primary and secondary schools in the Region, whose experience as a principal ranged from two to approximately 20 years and whose age ranged from early thirties to mid-sixties. All were known to one another.

There were also differences between the two groups. Group 1 principals had previously worked in the same district within the Region where the current regional director had been their district director. They liked and trusted him and had taken part in an informal and voluntary type of external review that he had initiated. They all volunteered to be in the Cyclical Review program because they were positively predisposed to the concept of Cyclical Reviews. They also held views about what could be improved, the most cited being the need for the school to receive a report as soon as possible after the review. These principals also negotiated (each with another Group 1 principal) to lead their own schools' Cyclical Reviews. The choice was theirs: one primary school principal negotiated with the principal in the adjacent high school. The two schools worked as a tightly coupled learning community, with almost all primary students feeding into the high school. A high school principal negotiated with another high school principal. Both knew and respected each other but had not previously worked closely together. Their schools had different socio-economic compositions and some differing programs. They were keen to learn from each other.

The Group 2 principals came from schools throughout the Region. Only two of the nine had had the current regional director as their district director but all were willing to trust his direction, although possibly not to same extent as the Group 1 principals. Their district directors recommended them as leaders who would benefit from or contribute well to a review as a team-member principal; again, being voluntary, the decision to participate was theirs. These principals were offered, but could decline, a review that appeared to be a good fit in the eyes of their director and the two Group 1 principals for the review. For example, one Group 2 principal with technology expertise joined a review team of a school whose principal wanted to expand this area of the curriculum.

To undertake the research, a theoretical model of evaluation influence (Ikin and McClenaghan 2015), designed to map the influences as they occurred, was developed from the relevant literature, including, in particular, the theories of evaluation influence proposed by Kirkhart (2000) and Mark and Henry (2004). This model combined Kirkhart's construct that influence interacted across a three-dimensional matrix with dimensions of source (results-based or process-based influences), intention (intended or unintended), and timeframe (when the influence occurred: immediate, end-of-cycle, or long-term) and Mark and Henry's extension of this theory, which included the mechanisms and outcomes of evaluation (general, cognitive and affective, motivational, and behavioural) that influence attitudes and actions at different levels of analysis (individual, interpersonal, and collective). In addition, to ensure that the data addressed the first and third research questions, the model allowed for the factors that triggered influence (based on Alkin's 1985 categories) and the outcomes, based on evaluation capacity building theories (Baizerman et al. 2002b) to be mapped separately. Thus the model was designed to identify separately the factors that triggered influence and then capture and map the strings of influence as they occurred, to try to determine which factors or influences were necessary precursors for further influences, to understand how long influences lasted, and to gauge whether principals built evaluation capacity and implemented other evaluation practices in their schools as a result of participation in the process (Fig. 10.3).

Data were collected from both groups at the three times (Immediate, End-of-cycle, Long-term), shown in Fig. 10.3, and in line with this model used for identifying themes, assigning codes and key words, sorting, and analysing. First, principals were observed before, during, and after each review. Second, principals completed three questionnaires: the first, immediately following the review, related to the factors they perceived to have either helped or hindered the review process; the second, completed a few months later, related to the immediate and medium-term influences they perceived; and the third, completed approximately 1 year after their participation, related to the long-term influences they perceived. Third, documents relating to the preparation for and implementation of Cyclical Reviews, as well as those from each of the participating principals' schools, were analysed to look for evidence of influence. Fourth, interviews were conducted with four of the participating principals to extend the conversation begun with the questionnaires and thereby deepen understandings of principals' perceptions.

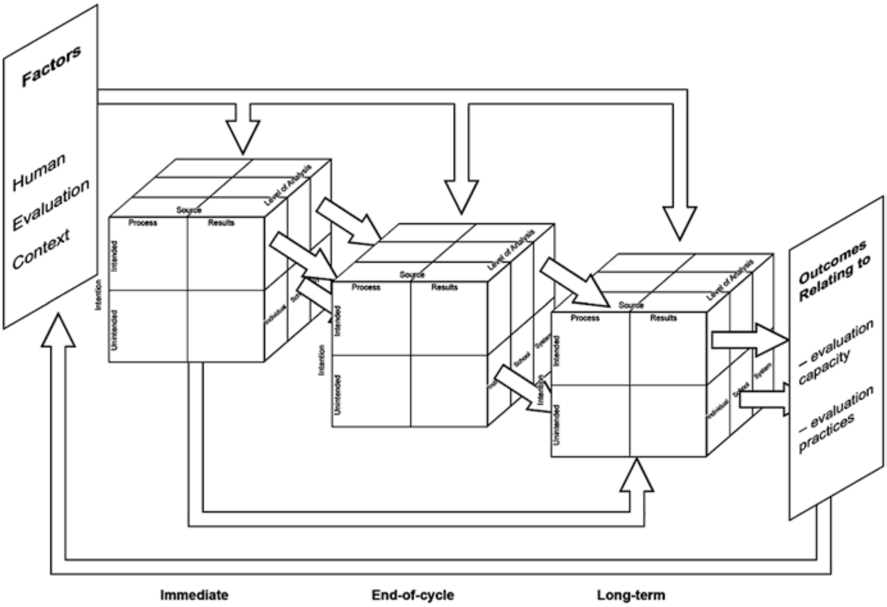

Factors
Human
Evaluation
Context
Source
Process
Results
Level of Analysis
Intention
Intended
Unintended
Individual
School
System
Source
Process
Results
Level of Analysis
Intention
Intended
Unintended
Individual
School
System
Source
Process
Results
Level of Analysis
Intention
Intended
Unintended
Individual
School
System
Outcomes
Relating to
– evaluation
capacity
– evaluation
practices
Immediate
End-of-cycle
Long-term

Fig. 10.3 Model for the analysis of evaluation influence (Kirkhart 2000, p. 8, and Mark and Henry 2004, p. 41)

10.5 What Changed?

The findings from the research project, described below, showed that this type of peer-review program, based on participatory, collaborative, and empowerment evaluation approaches, changed the traditional practice for school accountability and improvement reviews that were used in the Region, and influenced the principals to change their perceptions of and attitudes towards reviews and professional accountability, and the way review processes and results were used in their schools.

The findings also showed that there were a number of factors that could trigger these changes, and even more importantly that there were underlying values, such as trust, openness, and transparency that acted as catalysts for change. Further, the research showed that while this peer-review approach added significant value to all the participating principals and their schools, there were marked differences in the influences that participation had between the two groups of principals.

10.6 How Did Principals Perceive the Approach to Be Different from Traditional Practice?

The descriptions provided by the principals showed that the changes that they and their schools underwent and the subsequent long-term benefits were influenced by the role each Group played in the program and the evaluation approach used for this role.

For Group 1 principals an empowerment evaluation approach, based on its principles of improvement, community relationships, inclusion, democratic participation, social justice, community knowledge, evidence-based strategies, capacity building, organizational learning, and accountability, was employed from the outset, using the 10 steps from the Getting to Outcomes® (GTO) approach (Wandesman et al. 2000), but not necessarily in the same order. These steps and their significance are outlined in Table 10.2. Group 1 principals were the key stakeholders and were given the responsibility of designing the Cyclical Review program and implementing it as a pilot in their own schools. As the leaders not only of their own schools, but collectively within the community of participating schools, they made the decisions and were accountable for the outcomes. For these principals, empowerment evaluation changed how they participated.

> No-one made us develop Cyclical Reviews. The basic idea was run past us by John[2] [Regional Director] and Tara [critical friend]. We trusted them … John proposed that Tara work with us because she had expertise in evaluation. Tara helped us …, but as a coach, facilitator, critical friend—she wasn't the team leader. She told us what she thought, provided training, and gave us insights into evaluation practice, but it was up to us to make the decisions. It was our program. So, it was either going to work or fail because of us, and we were going to make sure that it worked. This approach just hasn't happened before, not for accountability anyway. (Group 1)[3]

Group 1 began by analysing the needs and necessary conditions for attaining the desired outcomes outlined in the regional director's brief (GTO, Step 1). The critical friend provided guidance on the types of tools and methods that could be used in this peer-review process, but did not tell them what to do. Based on the evidence that they had collected and discussed, the Group 1 principals iteratively developed the program's scope and sequence (GTO, Steps 2–7).

> The analysis and scope for developing the model and then implementing it made us conceptualize the whole program. There it was, in a logical model, and with all our thinking behind it—about why we were going to do it this way, what instruments we would use, who would be on the teams, how long it would take, how the reviews would run, and so on. That was the crux. We understood why we were going to do it this way. (Group 1)

---

[2] Individual's names have been changed to maintain anonymity.

[3] The quotes provide summaries of the collective voice identified as only Group 1 principals, Group 2 principals, or both Group 1 and Group 2 principals based on comprehensive and trustworthy data collection as described in the case-study section.

Table 10.2 Getting to outcomes for Group 1 principals

Step	Evaluation/accountability question	Significance of the step for Group 1
1. Needs assessment	"What are the needs and resources in our schools individually and as a community of schools?"	Empowered to assess needs and gaps
		Opinions valued; confidence and trust built
2. Goal setting	"What are the goals, target population, and desired outcomes (objectives) to fit the regional brief and for our schools individually and as a group?"	Ability (with assistance of critical friend) to set realistic objectives; define and delimit scope of program
		Ability to develop a logic model
3. Science and best practices	"How does the Cyclical Review program incorporate knowledge or science and best practices into our current school improvement and accountability practices?"	Oriented to evaluation processes
		Increased knowledge of wider evaluation research, theory, and approaches
		Empowerment evaluation process demystified
		Increased knowledge of evaluation best practices
4. Fit; culture context	"How does the Cyclical Review program fit with the review policy and practice already in place?"	Increased knowledge of current school review processes and policy
		Analysis of own school's planning and evaluation processes
5. Capacity	"What capacities are needed to put the Cyclical Review program into place with quality?"	Analysis and increased understanding of current Exemplary Practice Statements
		Increased knowledge of data collection strategies to be used and research to support their use
		Increased skills in designing and implementing the data collection tools and carrying out the Cyclical Review
6. Planning	"How will the Cyclical Review program be carried out?"	Ability to plan each school's Cyclical Review
7. Implementation/ process evaluation	"How will the quality of implementation be assessed?"	Ability to execute each school's Cyclical Review
		Ability as the Advisory Group to describe the processes and monitor the implementation

(continued)

Table 10.2 (continued)

Step	Evaluation/accountability question	Significance of the step for Group 1
8. Outcome and impact evaluation	"How well did the Cyclical Review program work?"	Increased use of data in school planning and evaluation
9. Total quality management; continuous quality improvement	"How will continuous quality improvement strategies be incorporated?"	Ability to critically analyse all aspects of program and make adjustments (double-loop learning)
		Used the reviews to develop a culture of organizational learning in their school
10. Sustainability and institutionalization	"If the Cyclical Review program is (or components are) successful, how will it be sustained?"	Principals became advocates for reviews to be adopted throughout the region and beyond
		Principals used these principles for other in-school evaluations

Cyclical Reviews were designed to evaluate how well a school was performing, to assist the school to build on its successes, and re-evaluate areas meriting attention (Fetterman 1994; Fetterman et al. 2015, 2018d). That is, they were designed to start from a positive position and build upon it. Reviews-by-exception, on the other hand, judged school performance deficits that were based on terms of reference which had been developed by an external expert. Cyclical Reviews created a challenge for the Group 1 principals, who, as a first step, had to develop statements about or standards of what an exemplary school would look like (GTO, Step 5). It made them think about evaluation as an integral part of their schools' planning processes, rather than as an inconvenient imposition that occurred when their annual school reports were due (GTO, Step 4).

> The Exemplary Practice Statements play a key role in the evaluation process. They are necessary to clarify what constitutes a good school from our point of view. Once this aspect is understood the rest of the process falls into place. At first, we thought that we would just use what the department had developed for us a long time ago. Most of us still used the pre-prepared surveys to fulfil our obligations for our annual reports. But Tara convinced us that we needed to make them ours. What we thought might be a short meeting turned into months of work, but it was worth it. They took on a new meaning. We wanted to see our school through these statements. We were aware that if Cyclical Reviews were going to last beyond the pilot, these statements needed to resonate with other principals across the Region. We developed them, but consulted with the regional principals' groups along the way. (Group 1)

The Group 1 principals believed that if the Cyclical Reviews were based on empowerment evaluation principles (Fetterman and Wandesman 2015), the process itself would provide invaluable professional learning, especially evaluation capacity building, for both themselves and individual team members, and would also facilitate organizational learning and deliver outcomes and accountability (GTO, Step 5).

> We want to be involved in order to further self-evaluation, evaluation, development, and involvement of our staff. We want confirmation of our successes and ideas on how to further them. But we also want to learn … what we can do better and ideas to think about in that regard. (Group 1)

Once the project had been approved at a regional level, the Group 1 principals developed all the components for the reviews (Fig. 10.1), with specific skills training when needed. For example, principals were used to administering surveys but not developing their own valid and reliable survey instruments. Skills training was thus provided by the critical friend in survey design, as well the development and use of the other evaluation instruments, such as developing interview questions, electronic data collection, coding of information, and the analysis and triangulation of evidence (GTO, Step, 5).

> Tara provided this training. Some of us only used our computers to read emails or write a letter, not for data entry. Our prior experience in school reviews was of the pen and paper type, sometimes butcher's paper. Using hand-held devices that could link to a spreadsheet on a computer and then sort the data took some getting used to. As the time progressed though we acquired a degree of expertise, because we were totally involved, we were doing the developing. Tara didn't do it for us. (Group 1)

Although Group 1 principals developed the evaluation tools in subgroups, they met regularly as an advisory group to discuss progress and challenges and to trial, assess, and refine the instruments. The critical friend attended every meeting of the advisory group and, when required, their subgroup workshops (GTO, Steps 9–10).

> After the first few reviews, we asked Tara to work again with us on [the desk audit]. We came up with a much better approach that worked really well—more a school self-evaluation. It now sets up the context. We understand the purpose of a desk audit much better now … it's an important factor in understanding evaluation and how to use data at school level because it means principals have to be self-critical and self-reflective. It's important because the principal can select the evidence. (Group 1)

Group 2 principals used a combination of participatory and collaborative evaluation approaches, and therefore although they became an integral part of a review team, they did not design the program, nor did they develop or refine the instruments, and, at least in the pilot phase, their schools were not reviewed. For them the program was collaborative: they were engaged as stakeholders; they formed part of the wider regional principals' groups, which were consulted during the development of the exemplary practice statements and kept informed about progress during the program's development.

> Once we realized that there would be other principals on the review teams, we were keen to take part and find out about it from the inside before committing to a review of our schools. (Group 2)

The program for Group 2 principals was also participatory to the extent that they shared control of the school review in which they were a team member—collecting and analysing data, reporting and disseminating results, and providing feedback to the advisory group on the process in which they had been involved for further development of the program. A combination of participation and collaboration opened a new avenue for their own professional learning, built new capacity in evaluation methods and practices, and for some provided the catalyst for their ongoing support of and involvement in the program. Because they did not develop the process, iteratively refine it, or see it applied in their own schools, they did not, however, report the same level of ownership or understanding as their Group 1 peers, although they did report enjoying the experience, gaining some ideas for their own schools, and learning some things about evaluation. Overall, Group 2 principals had a wider range of views about the change in approach and what they perceived to be different.

> Some people would see the Cyclical Reviews from a developmental perspective. Some from an accountability one. As it is voluntary, the [regional] principals' [associations] are comfortable and so see it as professional learning … development. We were looking forward to working with a team to see the mechanics of a sound evaluation process and learn from it. The process was a systematic way of collecting data. Better than methods in previous reviews, school development policy, Quality Assurance. There was a consistency in the way it was used to find themes and findings. For us at this stage it is developmental. If we take one on, then the focus will change. The hosts have a different perspective. They also wrote the instruments and procedures and feel they own it. But from a teacher's perspective it is accountability. These perceptions are important for later use of the results in the school. For us it is a mixture. (Group 2)

10.7 How Did the Reviews Influence the Principals?

There were process-based and results-based influences on both groups of principals; that is, principals were influenced by how reviews were conducted as well as by the findings and recommendations of the reviews themselves. Although there were some influences for individual principals, the research identified patterns of influence that emerged over the course of the reviews for each group. Principals reported that the influences that they experienced were a direct result of the process being one of peer review, rather than an imposed external evaluation. They attributed this to their ownership of the process and the trust placed in them. In contrast to the anecdotal evidence from previous externally imposed and run reviews, the impact did not seem to dissipate with time.

10.8 Results-Based Influences

Logically, the results-based influences were mainly reported by and observed in the Group 1 principals; it was expected that the findings and recommendations would be considered and acted upon in the reviewed schools. The Group 1 principals were the ones whose schools were reviewed. Interestingly, the influences they experienced strengthened over time and tended gradually to shift from influences on themselves as individuals, to influences on themselves as leaders in their schools; to influences on themselves as system leaders. Typically, principals explained that as individuals, they each needed to gain an in-depth understanding of the Exemplary Practice Statements, because the effectiveness and acceptance of the review results and processes hinged on a common understanding of these statements. They further explained that the review itself, especially the data and report, provided the opportunity for them to confirm or change their own opinions and attitudes about their schools and improve their knowledge of their staffs and schools. As leaders in their schools, they reported and data showed that they were becoming agents for change and advocates for the findings. They openly addressed and led discussions about the review data, findings, and report with their executives, staffs, and communities. They reported being more at ease talking in a range of forums about their schools' progress. As system leaders they spoke positively about this type of review in regional principal meetings and also in these forums commented on the legitimacy of the results of the reviews.

> It's not so much that we changed … it's more that our attitude towards things changed and we felt much more comfortable affirming what we now know is happening and why. (Group 1)

The most significant change in their thinking for Group 1 principals was that the changes resulting from the review were now the 'new normal' in their schools' cultures. It was envisaged, possibly naïvely, that principals would continue to make reference to the review when discussing, analysing, or evaluating school planning, strategies, and directions with their staffs, even after 1 or 2 years. This was not the case. Group 1 principals, although conscious of the ongoing influence of the review, specifically noted that the changes had now become embedded into school practice and culture.

> The results and directions seem embedded … into the school plan. The schools have moved beyond the need to reflect on their particular reviews. It's the [school] plans that have now become the direction of the schools. So the reviews are not referred to directly but the actions are. (Group 1)

There was one influence that emerged for both Group 1 and Group 2 principals. It was intended that all principals would gain valuable professional learning from being on a review team and take back ideas, initiatives, or strategies to their own schools for discussion and implementation, and this certainly occurred. What was a little surprising was that both groups of principals transferred and applied a specific finding from the school being reviewed to their own schools, without further discussion or consultation in their own schools. A typical response from both Group 1 and Group 2 principals in this regard was:

> I could see that our school was at the same point in our development regarding the use of technology as this one, and for the same reasons. I took this finding to be equally applicable for my school and took this back to the executive [for immediate implementation] for our reporting purposes. (In this case made by Ben, a Group 2 principal)

10.9 Process-Based Influences

Empowerment evaluation approaches by Group 1 and participatory–collaborative approaches by Group 2 led to long-term differences.

Again, patterns emerged over the course of the reviews for each group. These included substantial similarities in immediate influences on both groups: a marked divergence of influences at the end-of-cycle with this divergence's being sustained in the long term; the very large number and diverse nature of influences on Group 2 principals at all three phases compared with the more focused and analytical approach adopted by the Group 1 principals; and the gradual emergence of different system-leadership influences for both groups. Taken together, these patterns illustrate the marked difference in the degree of evaluation capacity achieved through the two peer-review approaches used by the two groups.

At first, both groups of principals displayed greater attention to evaluation practices, increased knowledge and skills about evaluation, attitudinal change towards evaluation design, motivation to learn, and demonstrated newly acquired skills.

> We enjoyed spending the time thinking about evaluations. You don't give yourself that sort of time usually. There are too many other things to do. This made us do it. We thought more deeply about the evaluation process, the what, why, when of evaluation. We all increased our skills—discussing the Exemplary Practice Statements, designing, planning, and implementing reviews, developing targets and strategies for improvement. We feel confident interviewing, observing, and so on. Most important we realized that we needed credible evidence, not just our intuition. Reviews must be based on data. (Group 1 and Group 2)

At the end-of-cycle the Group 2 principals still displayed very similar process-based influences compared with their immediate phase influences. They continued to use a wide range of personal theories about evaluation and tended to use processes on an ad hoc basis. The exception was that by the end-of-cycle there was no further evidence of their attention to evaluation design, but they were using evidence increasingly in their own schools, which may have been a legacy of the newly learnt skills.

> We now look for data to back up observations and look to see where data are coming from—that is, coming from more than the teachers. Our school evaluations are much more dependent on data and triangulation. We are much more aware of the need for this and check that executive and staff can back up their plans and outcomes. General staff meetings and minutes, for example, have agendas devoted to evaluation and data analysis. (Group 2)

In sharp contrast, by the end-of-cycle the Group 1 principals had progressed to embedding cultural change through the more logical and systematic processes promoted by their empowerment approach to Cyclical Reviews. They became far more analytical as a result of their early ownership of and their subsequent engagement in coherent review processes. At the same time, they were reporting changes of attitudes towards staff members on the review teams and more open and transparent practices. This divergence continued in the long term with few exceptions.

> Looking back, it is the process that makes the reviews worthwhile. We all agree we need to modify some components … make them easier to use … and step up some parts of the

> training, but the process makes the reviews work. The staffs accept the results. So did we, because they really confirmed what we knew was good or needed to be done. But they also help staff to open up, to discuss things more, to question more. To start with, we were happy to go along with the idea and be in the pilot. It's more than that now. It's a good process that we have developed and it helps us develop. It's not just what someone else wants or says … we should be aiming for. The reviews have had a positive impact on the way we do business—facts, data, listening to each other, openness, transparency. The Cyclical Review model is far superior to other models, it builds capacity. We have taken it back to other operations in our schools and applied the processes systematically. (Group 1)

Regarding system leadership, both groups advocated Cyclical Review processes beyond the school. This endured through the end-of-cycle phase, but in the long term Group 1 principals tended to advocate for Cyclical Reviews, while the Group 2 principals advocated for other principals to take part in Cyclical Reviews as team members. In each case they were advocating for others to have the experience that they had had. Both groups called for modification of some of the Cyclical Review processes and most Group 2 principals sought opportunities for staff to take part in reviews within and beyond their area. This can be taken as indicating further interest in directing their own and others' professional development.

10.10 Interaction Between the Results-Based and Process-Based Influences

While there were similarities in the nature of learning for principals in both groups, the empowerment evaluation approach, which included the start-up requirements for the program, required Group 1 principals to understand and apply criteria in the evaluation of their own schools, to design evaluation strategies and techniques, and later to reflect critically on the quality of evaluation practices and capacities to improve the values and assumptions influencing their own practices. That is, they were engaged in a kind of learning that fully integrated an experiential learning cycle of experiencing, reflecting, thinking, and acting. This meant that Group 1 principals were actively engaged in designing their own learning, and this learning was further deepened by leading the peer-review team and gaining feedback by participating in the peer-review of their own schools. This is consistent with double-loop learning (Argyris 1996; Argyris and Schön 1978): learning that requires a shift in the mental model on which the decision depends.

In contrast, the Group 2 principals, through their participatory–collaborative approach, only had to commit to participating in a school-review team, without having the challenges and advantages of gaining empirical feedback about the effectiveness of their understandings of school evaluation. It was found that the principals who advised only as team members in other principals' schools were concerned with problem-solving mainly when outcomes fell short of objectives. This is consistent with the limits of single-loop learning: learning by repeatedly attempting the same problem with no variation or ad hoc variations of method and not questioning the goal (Argyris 1996; Argyris and Schön 1978).

10.11 The Factors Underlying the Influence of Cyclical Reviews on the Participating Principals

Regardless of the influences of the reviews, the factors that triggered these influences and the underlying values held by the principals with regards to each of these factors acted as catalysts for change, and they acted in a similar way, although to different degrees, for both groups.

Eleven characteristics that triggered the influences experienced by the principals emerged from the data. These characteristics were not dissimilar from those identified in the literature (Cousins and Leithwood 1986; Johnson et al. 2009) and fitted best into the three categories—context, human, and evaluation factors—proposed by Alkin (1985).

The context factor included such characteristics as the culture of the school being reviewed, the requirements of the school regarding the focus of the review, and requirements and the restraints of the review itself. Principals from both groups maintained that an understanding of the school’s culture and climate was essential for an accurate report and for the wider school staff, not just the principal, to have confidence in the findings. They also understood the importance of skills training for all team members, with Group 1 principals recommending that the training be accredited at various levels, such as team member, team leader, critical friend. They further recognized that when teams had a good balance of experience, expertise, and background across the schools’ learning communities, objectivity was enhanced and findings were more astute. Regarding duration:

> Three days were fine—although action-packed, essential, and time to provide a snapshot. We think any more than that would be too much … too much for the school, too much for other principals to be away from their schools. We felt a great sense of achievement to report on the schools in such a comprehensive fashion. It was the right focus for this type of evaluation. (Group 1 and Group 2)

The human factor included such characteristics as the credibility of the principal as an evaluator—principals identified expertise and interpersonal skills as essential. They noted that positive relationships were crucial, commenting, ‘It is difficult to effect change based on evaluation of sensitive data if the leader, host, and facilitator do not have trust in one another’ (Group 1 and Group 2). Commitment by the principal of the reviewed school to use the evaluation findings was another characteristic that triggered influence. Commitment was achieved through collaborative choice of team leader, and appropriate and timely input from the critical friend. Other characteristics included the motivation by both groups of principals to be involved, the interaction between the leadership of the review teams, the leadership by host principals, and the positive human relationships within the team and the rest of the school.

The evaluation factor included such characteristics as the structures and resources developed for the reviews. Both Group 1 and Group 2 principals believed that having all the tools—the survey, interview questions, desk audit, classroom observation list, and report format—tied to the Exemplary Practice Statements was critical. While all agreed that some of these had teething problems, it was the Group 1 principals who further suggested that further modifications were needed for some of these problems. Data collection processes and methods developed for the reviews also had a positive influence.

> Once we got the hang of coding, data entry, and sorting, the process proved a great way to go. When coupled with the data-point entries and survey results, the report provides the school with a comprehensive set of data for further investigation and future directions. (Group 1 and Group 2)

While the Cyclical Review evaluation methodology is primarily independent as a process, it was affected by varying school contexts and human factors, as described above. The major variance to standardised review processes appeared to be the leadership services provided to the review team, the host principal, and the partnership between the review leader and the critical friend. The principals' comments and actions showed that the evaluation capacity that was built was also dependent on whether their participation was through an empowerment process or a participatory–collaborative one. Ultimately, the Group 1 principals' evaluation capacity grew extensively compared with that of Group 2.

Finally, it emerged in the discourse when the principals theorized about the above-mentioned factors that their personal core values, for example trust in others, the need for transparency in processes, underlay the degree to which these factors triggered influence (Table 10.3). It can be seen that factors listed in Table 10.3 sit largely under the umbrella of trust as defined by (Tschannen-Moran 2014). This is a significant observation as Tschannen-Moran's and her colleagues' research (Goddard et al. 2001; Hoy and Tschannen-Moran 2007) showed that trust improves student performance and schools.

What can be asserted as a result of this research is that early and deep learning about evaluation processes gained and sustained through an empowerment

Table 10.3 Factors that triggered influence

Human factors	Context factors	Evaluation factors
Openness	Openness	Openness
Trust	Readiness	Clarity
Credibility	Clarity	Consistency
Competence	Transparency	Standards
Knowledge		Consensus
Commitment		Engagement
Ownership		Transparency
		Functionality

evaluation approach was essential to achieving the full benefits of a peer-review process. Group 1 principals, engaging in deep learning, created fresh structures in their schools and fresh outcomes, which illustrates Giddens's (1984) theory about the duality of structure: it is both process and outcome. In this case the Cyclical Review processes delivered new structures in schools and new professional leadership capacities in evaluation.

10.12 Future Directions

As a result of the pilot program, the regional director incorporated the Cyclical Review program into the regional strategic plan, so that it was available to all schools in the Region, but still on a voluntary basis. Principals were initially interested but slow to commit. Nevertheless, each year for the next 6 years more regional schools committed to a review and more principals asked to join the program. A steering committee was established to oversee implementation; the advisory group continued work on the tools and processes; some Group 2 principals became team leaders in the next round of reviews; and a few Group 1 principals became critical friends. The Group 1 principals and some Group 2 principals used a modified form of the empowerment approach to Cyclical Reviews within their own schools.

At a State level (influenced by the success of the Cyclical Review program) the idea of peer review gained momentum and a mandated peer-led cyclical review was mooted. Personnel and organizational changes at senior executive levels, however, brought different ideas to the fore. What finally eventuated in 2015 was a new planning and evaluation model, which also marked the beginning of a new era in school accountability and improvement for the New South Wales government school system. The model introduced an integrated school self-evaluation, planning, and evaluation process. A school's plan would be endorsed by the principal's supervisor as developed and completed in accordance with policy, while ultimate responsibility for the plan, its execution, and evaluation would rest with the principal. In addition, a small team of principal peers would assess the school's plan and achievements through an external school validation process.

This new model presumed a high level of competence in collaborative strategic planning and evaluation as well as a high level of evaluation capacity by all school principals and staff from the outset. It also appeared to presume that peer 'experts', selected on merit as excellent school leaders, although not necessarily on their evaluation expertise, would provide the 'trust' that had proved so necessary from the Cyclical Review program. The research findings described in this chapter provided little support for these presumptions.

This new model is now in operation across the State, and the system is providing some professional development for principals in evaluation. Nevertheless, a new approach is required to build evaluation capacity at a school level. Recent work by independent researchers, although on a small scale, is showing promise at the school level (McClenaghan and Ikin 2017). Based on the Fetterman's 3-step empowerment evaluation approach (Fetterman 2015), the schools in the independent research have engaged a critical friend to work with them as they develop and evaluate their school plans. These schools report not only being well prepared for the external validation, but also now having all staff owning and understanding their plan and how to reach their goals.

10.13 Conclusion

Fetterman et al. (2018a) suggest that although each stakeholder involvement approach can and often should be used by itself, the evaluation field is prepared for a new era of approaches to evaluation experimentation, with one promising path being a combination of approaches within the same initiative. This peer-review program was possibly a little ahead of its time, using a combination of empowerment and participatory–collaborative evaluation processes; empowerment for one group of principals, and participatory–collaborative for the other. Both changed traditional evaluation roles, changed the influence that a review could have on school leaders, and changed the impact of school reviews. The greatest impact, however, was felt by the principals engaged in an empowerment evaluation, particularly in the degree to which evaluation capacity was built.

References

Alkin, M. C. (1985). *A guide for evaluation decision makers*. Beverly Hills: Sage.

Argyris, C. (1996). *On organizational learning*. Cambridge, MA: Blackwell.

Argyris, C., & Schön, D. (1978). *Organizational learning: A theory of action perspective*. Reading: Addison-Wesley.

Baizerman, M., Compton, D. W., & Stockdill, S. H. (2002a). Editors' note. New directions for evaluation. In *The art, craft, and science of evaluation capacity building* (Vol. 93, pp. 1–6). San Francisco: Jossey-Bass.

Baizerman, M., Compton, D. W., & Stockdill, S. H. (2002b). New directions for ECB. *New Directions for Evaluation, 2002*(93), 109–120. https://doi.org/10.1002/ev.45.

Cousins, J. B., & Leithwood, K. A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research, 56*(3), 331–364.

Fetterman, D. M. (1994). Empowerment evaluation. *Evaluation Practice, 15*(1), 1–15.

Fetterman, D. M. (2015). Empowerment evaluation: Theories, principles, concepts, and steps. In D. M. Fetterman, S. J. Kaftarian, & A. Wandersman (Eds.), *Empowerment evaluation* (2nd ed., pp. 20–42). Thousand Oaks: Sage.

Fetterman, D. M., & Wandesman, A. (Eds.). (2015). *Empowerment evaluation principles in practice*. New York: The Guildford Press.

Fetterman, D. M., Kaftarian, S. J., & Wandersman, A. (Eds.). (2015). *Empowerment evaluation: Knowledge and tools for self-assessment, evaluation capacity building, and accountability* (2nd ed.). Thousand Oaks: Sage.

Fetterman, D. M., Rodriguez-Campos, L., Wandersman, A., O'Sullivan, R. G., & Zukoski, A. P. (2018a). Conclusion: Highlighting the present and looking to the future. In D. M. Fetterman, L. Rodriguez-Campos, & A. P. Zukoski (Eds.), *Collaborative, participatory, and empowerment evaluation: Stakeholder involvement approaches* (pp. 135–144). New York: The Guildford Press.

Fetterman, D. M., Rodriguez-Campos, L., Wandersman, A., O'Sullivan, R. G., & Zukoski, A. P. (2018b). An introduction to collaborative, participatory, and empowerment evaluation approaches. In D. M. Fetterman, L. Rodriguez-Campos, & A. P. Zukoski (Eds.), *Collaborative, participatory, and empowerment evaluation: Stakeholder involvement approaches* (pp. 1–9). New York: The Guildford Press.

Fetterman, D. M., Rodriguez-Campos, L., Wandersman, A., O'Sullivan, R. G., & Zukoski, A. P. (2018c). Similarities across the three apporaches: Principles and practices in common. In D. M. Fetterman, L. Rodriguez-Campos, & A. P. Zukoski (Eds.), *Collaborative, participatory, and empowerment evaluation: Stakeholder involvement approaches* (pp. 118–134). New York: The Guildford Press.

Fetterman, D. M., Rodríguez-Campos, L., & Zukoski, A. P. (Eds.). (2018d). *Collaborative, participatory, and empowerment evaluation: Stakeholder involvement approaches*. New York: Guilford Publications.

Giddens, A. (1984). *The constitution of society: Outline of a theory of structuration*. Berkeley/Los Angeles: University of California Press.

Goddard, R. D., Tschannen-Moran, M., & Hoy, W. K. (2001). A multilevel examination of the distribution and effects of teacher Trust in Students and Parents in urban elementary schools. *The Elementary School Journal, 102*(1), 3–17. https://doi.org/10.1086/499690.

Hoy, W. K., & Tschannen-Moran, M. (2007). The conceptualization and measurement of faculty trust in schools. In *Essential ideas for the reform of American Schools* (pp. 87–114). Charlotte: Information Age Publishing.

Ikin, K. B., & McClenaghan, P. (2015). Modelling the influences of evaluation on school principals: Towards evaluation capacity building. *Evaluation Journal of Australasia, 15*(1), 19–27.

Johnson, K., Greenseid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on evaluation use: A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation, 30*(3), 377–410.

Kirkhart, K. (2000). Reconceptualizing evaluation use: An integrated theory of influence. In V. J. Caracelli & H. Preskhill (Eds.), *The expanding scope of evaluation use. New directions for evaluation* (Vol. 88, pp. 5–24). San Francisco: Jossey-Bass.

Mark, M. M., & Henry, G. T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation, 10*(1), 35–57. Retrieved from http://evi.sagepub.com/cgi/content/abstract/10/1/35. https://doi.org/10.1177/1356389004042326.

McClenaghan, P., & Ikin, K. (2017). *Leading empowered evaluations to develop trust and improve learning: Insights from qualitative research* (Research Conference Paper). Retrieved 1 Ocotber 2018, from ACER https://research.acer.edu.au/research_conference/RC2017/28august/9/

New South Wales Department of Education and Training. (2004). *School development policy: New South Wales Department of Education and Training*.

New South Wales Education Portfolio. (1990). *School-centred education: Building a more responsive state school system (Scott Report)*. Milsons Point: Management Review: New South Wales Education Portfolio.

Tschannen-Moran, M. (2014). *Trust matters: Leadership for successful schools* (2nd ed.). San Francisco: Jossy-Bass.

Wandesman, A., Imm, P., Chinman, M., & Kaftarian, S. J. (2000). Getting to outcomes: A resulrs-based approach to accountability. *Evaluation and Program Planning, 23*, 389–395.

Yin, R. K. (2003). *Case study research: Design and methods* (Vol. 5, 3rd ed.). Thousand Oaks: Sage.

Kerrie Ikin began her career as a teacher and executive member in government schools across New South Wales, Australia. For over 25 years she worked at the systems level of the education department as a director of schools, senior policy adviser, and senior administrator. As the director responsible for 32 schools in Sydney's north-west, Kerrie was responsible for the leadership development of those schools' principals and aspiring principals as well as the broader role of developing and managing evaluation and accountability structures and processes for the region's 240 schools. Kerrie currently works as a consultant and researcher in leadership, organisational development, and evaluation-capacity-building, and as a reviewer of major government school infrastructure. She is also an adjunct associate professor in the School of Education at the University of New England.

Chapter 11
Research-Informed Peer Review

David Godfrey and Karen Spence-Thomas

Abstract This chapter aims to outline a process of school collaboration with peer review at its heart, conceived at UCL Institute of Education by David Godfrey and piloted by the authors for the first time in 2016/17 in London. We outline some of the learning that has resulted for participating school staff and reflect on what we have learned as facilitators and evaluators of Research-informed Peer Review (RiPR). The chapter looks at the principles and theories that underpin RiPR and its relationship to other research-practice models. We present findings from follow up surveys two years since completing the first RiPR cycle, and apply additional learning from its implementation and adaptation in Chile (see also Chap.7). Finally we look at how evaluation theory is linked to evaluation policy in this model and on the potential impact of RiPR on teacher collective efficacy

11.1 Introduction and Aims of the Chapter

This chapter aims to outline a process of school collaboration with peer review at its heart, conceived at UCL Institute of Education by David Godfrey and piloted by the authors for the first time in 2016/17 in London. We outline some of the learning that has resulted from the process for participating school staff, reflect on what we have learned as facilitators and evaluators of RiPR clusters and attempt to outline the ways in which RiPR is designed to lead to school improvement and professional learning.

Below we introduce the context in which the model was developed and implemented for the first time in London. This is followed by a description of the processes and practices of Research-Informed Peer Review (RiPR), including some of its modifications over time. The chapter then looks at the principles that underpin RiPR; some of this has been adapted from a previous publication in which RiPR is compared to another research-practice innovation, Research Learning Communities (Godfrey and Brown 2019b). In this chapter, we are able to expand, incorporating

---

D. Godfrey (✉) · K. Spence-Thomas
UCL Institute of Education, London, UK
e-mail: david.godfrey@ucl.ac.uk

© Springer Nature Switzerland AG 2020

D. Godfrey (ed.), *School Peer Review for Educational Improvement and Accountability*, Accountability and Educational Improvement,
https://doi.org/10.1007/978-3-030-48130-8_11

further data from follow up surveys 2 years since completing the first RiPR cycle, and applying additional learning from its implementation and adaptation in Chile (see also Chap. 7). We also develop two further lines of conceptual thinking on linking evaluation theory to evaluation policy and on the potential impact of RiPR on teacher collective efficacy.

11.2 Background

RiPR began in 2016, with a pilot study involving 6 primary schools in London. The first year, working with a group of enthusiastic and confident schools, proved an excellent experience for us as facilitators, and also gave us valuable proof of concept. We have subsequently refined the model and have run clusters from the London Centre for Leadership in Learning as a knowledge transfer project with different groups of schools each year. The RiPR system has been taught to groups of principals in Bulgaria, Chile and Colombia. The initiative in Chile has also been set up as a research project with nine primary schools in three regions, working in collaboration with a local leadership centre, who have adapted the model to fit local context, renaming it Schools Inquiring and Learning with Peers (SILP) (see Chap. 7). The latter is planned to continue into a second year, and will be built into a larger, funded research project.

In England, the educational context is a dynamic one. Schools work in a highly competitive market; parents are able to preferentially rank choices for schools they wish their children to attend, there are school rankings and league tables of examination results published in national newspapers and a high stakes inspection regime with publicly accessible reports. Nevertheless, the picture is also one of greater collaboration; schools increasingly work in networks or alliances, such as through formal Multi-Academy Trusts (MATs) or Teaching Schools Alliances (TSAs) and many other informal local arrangements. Given the lesser uptake of academisation for primary schools, and the smaller scale of these schools, primary schools tend to be hungry for other local arrangements for collaboration in order to share resources or combine forces to improve aspects of educational provision. This has been particularly important in an era where Local Authorities (LAs) have lost significant portions of funding due to the academisation programme (where schools apply to be independent entities, directly funded by the state rather than operating under the auspices of LAs).

This environment of 'coopetition' (Muijs and Rumyantseva 2014) has created a demand for robust and rigorous processes for schools to work together while keeping a competitive advantage over their local 'rivals'. Thus, arrangements that allow schools to form their own strategic and voluntary alliances (such as RiPR) fit in well to this ecosystem (Godfrey and Brown 2019b). Coupled with a push from central government to engage in more evidence-based practices (Godfrey 2017), schools are eager to find structures that use research-informed principles to improve student learning and to lead pedagogical change. The use of school peer review among primary schools is also now commonplace in England (e.g. Greany and Higham 2018 and ch. 1 within). These peer reviews can provide a level of professional challenge that school leaders wish to provide to themselves and their leadership teams, often in anticipation of external inspections (Matthews and Headon 2015 and ch. 4 within). The RiPR model, while capitalizing on this eagerness to engage in peer review models, also provided a distinct offer to our schools, some of whom were already quite research-engaged and wished to engaged in a process that was less closely linked to an external inspection framework and more closely aligned to improving an aspect of teaching and learning. Thus, RiPR was seen as an innovative model, one linked to research and with the bonus of having a reputable university behind it. The model was emphasised as a collaborative, enquiry-focused approach to improving educational provision and contrasted with more 'summative' evaluation or 'Mocksted' approaches (i.e. rehearsal for an Ofsted inspection).

The RiPR model built on the shared interest of the authors and others in our Centre to use research-informed leadership processes to empower practitioners to improve their schools and to implement meaningful changes within them. Paramount in this process is to keep a focus on the desired impact – usually this is ultimately to do with – richly defined – improvements to student outcomes. The process and principles of RiPR are outlined in more detail below.

11.3 The Research-Informed Peer Review Process

In the RiPR model, schools work in clusters of three conducting reciprocal review visits during the school year. RiPR members comprise one senior leader (usually the Headteacher) plus at least one other teacher, middle or aspiring school leader, up to a maximum of four people per school. Participating staff also attend three half-day workshops, two before the first review visit and one at the end of the school year and after all three visits have been carried out. These workshops sometimes combine with other clusters. See Fig. 11.1 below for a visual outline of the process.

Workshop one begins with a discussion of a literature review sent out as pre-reading on a topic of educational interest and relevant to the cluster's improvement aims. Although by no means limited to this, so far RiPR groups have tended to focus on the topic of effective teacher feedback and assessment. These themes have wide appeal, not least because effective 'feedback' is one of Hattie's 'big-hitters' for improvements to student attainment (see Donohoo et al. 2018). It also makes more sense to examine teacher feedback use in the context of the schools' assessments policies and practices, in particular looking for the balance between formative and summative assessment (Black and Wiliam 2005). Thus, improvements in this area are anticipated to have a strong likelihood of achieving measurable gains in student test data. More widely, by focusing on teacher feedback and the school's assessment policies, the focus invariably leads to wider- and deeper – questions about the school's vision and understanding of (excellence in) student learning.

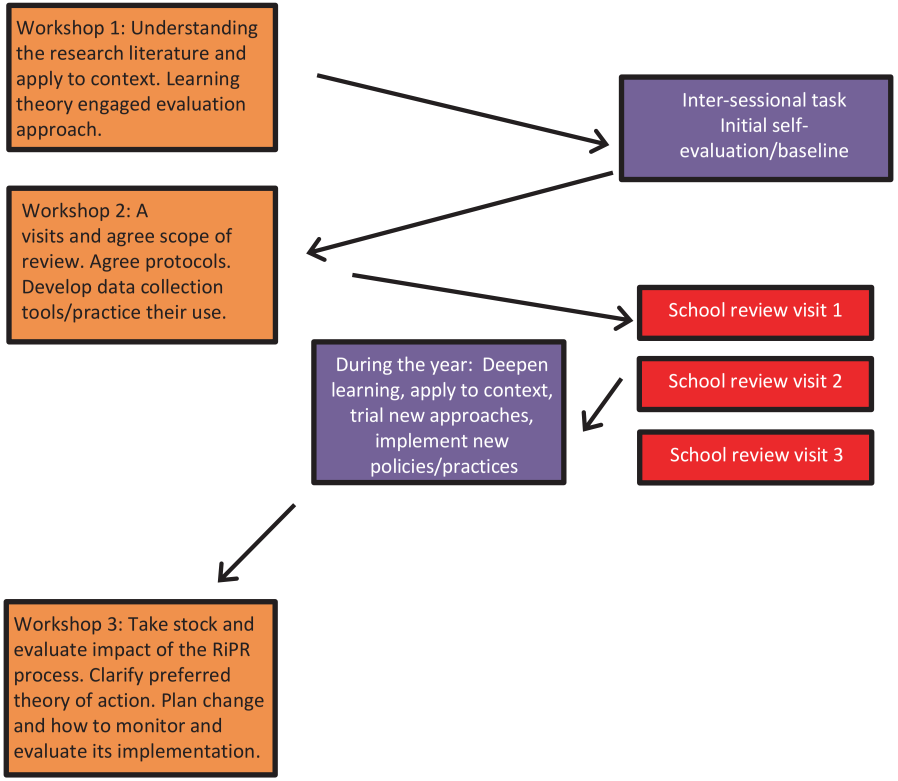

Workshop 1: Understanding the research literature and apply to context. Learning theory engaged evaluation approach.
Inter-sessional task Initial self-evaluation/baseline
Workshop 2: A visits and agree scope of review. Agree protocols. Develop data collection tools/practice their use.
School review visit 1
School review visit 2
School review visit 3
During the year: Deepen learning, apply to context, trial new approaches, implement new policies/practices
Workshop 3: Take stock and evaluate impact of the RiPR process. Clarify preferred theory of action. Plan change and how to monitor and evaluate its implementation.

Fig. 11.1 An outline of the research-informed peer review process

The other key feature of workshop one is the introduction of the notion of theory-engaged evaluation (Robinson and Timperley 2013) and, as part of this, the concept of theories of action. The former principle underpins the RiPR evaluation process and the latter concept informs participants' understandings of past, present and desired policies and practices for their schools. An important exercise in this workshop is the construction of a school's 'theory of action' for assessment at organisational level. This helps participants make sense of their current situation and to begin to see the strengths and weaknesses of their policies in both theory and practice.

For the second workshop, participants bring in their school's assessment and feedback polices and these are exchanged between schools for mutual scrutiny. Participants are guided to de-construct the theories of action inherent in these policies, in particular the underlying values and beliefs implicit in them, the actions they recommend or prescribe, and the outcomes they intend to produce. Among the three schools in the cluster and ourselves as facilitators, we agree a cycle of review visits, deciding on who will go first and what the scope of this first visit will be. We also guide participants to draw up a timetable and recommend how to get the most out of this, mixing classroom observation with interviews of staff, students and book scrutiny. We recommend protocols for the conduct of evaluation teams during the visit and we also share or co-construct data collection instruments relevant to the focus.

Each school visit lasts one whole day and is conducted as a process of enquiry and mutual learning, facilitated by a university representative who is experienced in the process. After a day of collecting and reviewing evidence there is a final meeting that follows a process of facilitated discussion to draw out the key learning from the day. Key findings and messages are summarised by the facilitator and sent to all participants.

After the three review visits, all participants in the cluster attend the final workshop in which the learning is consolidated. School teams reflect on their initial theories of action, share how these have developed over the course of the process and determine their preferred theories of action for the coming year. Participants evaluate how each of the seven key ingredients of the RiPR process have influenced their learning and impacted on their decisions, i.e.:

1. The shared enquiry theme (e.g. feedback/assessment)
2. The literature review (on the above theme)
3. School review visits
4. Co-design/use of data collection tools bespoke to the area of focus and the review
5. The principles of theory-engaged evaluation
6. External facilitation
7. Feedback from peers

School leaders are also shown how to plan for on-going implementation of their new improvement plans based on the Concerns-Based Adoption Model (Hall 2013). To sustain network-level working, schools are encouraged to continue to support each other and to build a cycle of feedback loops to inform future actions.

11.4 The Principles of Research-Informed Peer Review

The RiPR model shares features with joint practice development, collaborative enquiry, networked learning communities, professional learning communities and research learning communities. Elsewhere, Godfrey and Brown (2019b) have explained in detail the key principles that RiPR shares with another model as a research-practice approach. The explicit framework for evaluation and the school visits also lend a distinct flavour compared to the above approaches, and provide a particular focus for RiPR work which will be expanded below. The focus on theories of action in this model aligns particularly well with our Centre's approach to 'impact' in bespoke leadership development programmes with schools, in particular how we tend to start with the end in mind, and build a rich picture of how we visualize this future as different (and better) to current practices, (see Earley and Porritt 2014). It also follows an academic understanding of research-engagement that is quite distinct from the so-called 'what works approach' giving emphasis to combining different types of evidence and knowledge (both professional and academic) when coming to conclusions (see Godfrey 2017; Biesta 2007).

In this section we outline these features and develop some further issues of conceptual interest, particularly the relationship of school policy to theory and secondly the potential of peer review (especially this model) to promote collective teacher efficacy.

In '*An ecosystem for research-engaged schools: reforming education through research'* (Godfrey and Brown 2019a), the authors outlined a conceptual framework to understand the challenges of infusing research into the practices and policies of school education at every level. Central to the ecosystem model was the notion of 'research-engaged schools', such schools:

1. *Promote practitioner research among staff (especially teachers)*
2. *Encourage staff to read and make sense of published research*
3. *Welcome participation in research projects led by outside organisations such as universities*
4. *Use research to inform decision-making at every level of the school – individual, departmental, whole school and in collaborative work*
5. *Have an outward looking orientation, which may be aided by maintaining research-based links with other schools, universities or professional/academic entities.* (Handscomb and MacBeath 2003; Sharp et al. 2005; Wilkins 2011, adapted from Godfrey and Handscomb 2019 p. 9)

In the RiPR model we can see all of the above elements in practice, and thus the promotion of organisational learning and improvement at individual, organisational and network level are actively promoted in this approach. Participants work collaboratively to learn professionally, using academic research alongside practitioner enquiry and other data to engender changes to practice that ultimately are hoped to improve students' learning. These come about with the help of a university based facilitator or facilitators who sometimes adopt a dual role as researchers of the process. The latter point is also important when seeing RiPR as a 'bridge' between the worlds of academia and practice. Not only does the research provide greater incentive to academics to work alongside schools but the learning from this is then fed back into improvements in the model that ultimately help the schools who benefit from working with more effective school improvement and professional learning approaches.

The RiPR model promotes a model of research-to-practice that is quite different from the somewhat dominant 'what works' approach (see Godfrey 2017 for further elaboration of this argument). Thus, the specific intention of RiPR is to promote 'research-informed practice', i.e.:

> "An actively enquiring mode of professionalism that involves critical reflection and engagement in ('doing') and with ('using') academic and practitioner forms of research, taking into account both the findings and theories generated from them" (Godfrey 2017, p. 438).

The above definition makes explicit our modus operandi as academics and knowledge exchange professionals at a university working with practitioners. Not only does this explain the intention to combine different types of evidence and knowledge from academic research with that generated in the context of practice but also the role of teachers and school leaders in actively discovering and creating their own practice-relevant knowledge. Godfrey and Brown (2019b, p. 91–107) describe eight key features of this kind of research-practice model:

1. Schools engage in deep self-evaluation and enrich their understanding of their baseline in relation to a shared theme
2. Participants engage in reflective and collaborative professional learning and model this with their own staff
3. Practice-based knowledge is combined with research-based knowledge to create new knowledge
4. Key people participate in and lead change
5. Theories of action are made explicit when proposing improvements
6. External facilitation plays a crucial role
7. Change takes time and involves cycles of reflection and the development of knowledge and skills
8. Trust is essential and success is made more likely when schools support and challenge each other

We adapt these features of effective research-to-practice models in our description below, focusing on the principles of the RiPR approach:

1. This first principle has been highly valued by participants in RiPR in our evaluations.

One pilot school said the 'shared theme encouraged collaboration and allowed for deeper learning and thinking as the cycle progressed'.

The rationale for a shared theme across a cluster, as opposed to each school setting its own aims and priorities, has two overarching advantages. First, we are able to focus, in advance of the first workshop, on synthesizing what the academic research has to say about this issue, thus avoiding 're-inventing the wheel' or continuing to act on sometimes erroneous or limited prior knowledge of the issue. As part of the RiPR process we provide a bespoke summary of key literature on the topic (e.g. about effective feedback and assessment) and school participants are sent this prior to the first cluster meeting. Secondly, when sharing a theme of practice, participants deepen their understanding of the principles they derive from research, through successive iterations of comparison of learning derived from the literature review with classroom observations of teaching practice and other school data collected from each of the three schools. To illustrate this, we show a slide of archeologists uncovering an ancient site and gradually developing their understanding of it; under the slide we quote Fielding et al. (2005) whose mantra: "*digging deeper, digging together",* neatly summarises the principle of joint practice development that we encourage in the RiPR collaboration. The RiPR group is, at heart, a Professional Learning Community (PLC) and, as such:

> *"PLCs are most effective when they have a shared vision and sense of purpose; members of the PLC consistently taking collective responsibility for student learning; participants collaborating in ways that go beyond mere superficial exchanges of help or support; the promotion of group as well as individual learning; and participants engaging in reflective professional inquiry"* (Stoll et al. 2006, taken from Godfrey and Brown 2019a p. 96).

Given this particular emphasis in the RiPR model, it may be most aptly described as a form of 'collaborative peer enquiry', towards the more constructivist end of the peer review spectrum (see Chap. 13 for full analysis).

2. As facilitators of the RiPR approach, we hope to promote the PLC model of working so that the constructivist model of professional learning in the workshops and peer review visits are also mirrored in the opportunities that the school leaders provide to their own teaching staffs. Pilot schools in particular greatly appreciated the opportunity for structured peer-to-peer dialogue and feedback, one commenting their learning conversations had been 'focused, evidence-based, supportive, appreciative yet challenging'.

This is important, since for new ideas to be implemented effectively teachers need to have ownership of the change (Hall 2013), A pilot school commented that, "empowering staff is key to implementing change. Allowing staff to take risks within key parameters changes practice more quickly" (New Ash Green Primary School). This avoids the so-called 'bypass' approach so often employed by school leaders characterised by attempting to impose ideas on staff through persuasion rather than buy-in (Robinson 2017).

3. Participants in the RiPR process read the academic research on the shared enquiry theme and also receive extensive input into the literature on theory-engaged evaluation (Robinson and Timperley 2013) theories of action (Robinson 2017) and the process of implementation of innovations (Hall 2013). This intention to combine the explicit knowledge from academia with the tacit knowledge developed in the context of practice is redolent of the knowledge transformation model outlined in Nonaka and Takeushi's (1995) work on organisational learning.

Adding the collaborative and enquiry focus of RiPR we have a 'knowledge creation' model in RiPR:

> *"where practically useful and contextually pertinent knowledge results when the producers and users of formal knowledge, who are, simultaneously, also the users and holders of 'practical' knowledge, come together share what each group know"* (Godfrey and Brown 2019b, p.96).

Outcomes are improved through the development of new practices, policies or curricula changes in relation to the shared theme but broader, incidental learning always occurs as participants reassess their educational values.

4. In RiPR collaborations, we invite the participation of formal school senior leaders but also others who may be closer to the teaching practices. Thus each school sends at least two participants and sometimes up to four. Central to our thinking is the distinction between the strategic formal role of leaders (or managers) in the school and the informal leadership processes needed to successfully implement changes in organisations. The involvement of headteachers not only sends an important message to other staff that the process is valued and worthy of everyone's time and efforts – modeling the 'lead learner' approach (James et al. 2007), it also means that any organisational changes that might be needed to allow teachers to collaborate with others in the learning or changes that are planned, can be formalised by the headteacher.

Whilst the senior and other leaders involved in RiPR have formal roles at the school and this is important for the reasons stated above, leadership should also be seen as an informal process of influencing others, the latter often determining the success of new initiatives Spillane et al. (2010). Aspiring leaders or teachers known to be innovative and excellent practitioners may have the kind of high Practice-Based Social Capital (PBSC) described in the research (e.g. Baker-Doyle and Yoon 2010) and who can help persuade others to adopt new ideas and to spread the word around the school. Involving teachers in the peer review process is also great experience for more junior staff and forms a pedagogical bridge between teachers and school leaders, sometimes lacking in other peer review models.

5. In RiPR, theories of action are made explicit in a number of ways. First, the evaluation model itself is outlined to participants and this explains how the evaluation is proposed to lead to learning, followed by improvements to practice. For this, we use Robinson's and Timperley (2013) model of evaluation through theory-engagement, shown in Fig. 11.2 below:

Central to the RiPR evaluation model is an understanding of the distinction between practitioner theories (P-theory) and evaluator theories (E-theory). E-theory is informed by a reading of the relevant academic literature on a shared theme as well as through the additional data gathered from review visits to each school and learning from facilitated workshops. P-theories are those evident in the policies and practices at each school. Here we make a further distinction between ***espoused theories (the talk)*** and ***theories-in-use (the walk)***. While the former are inferred

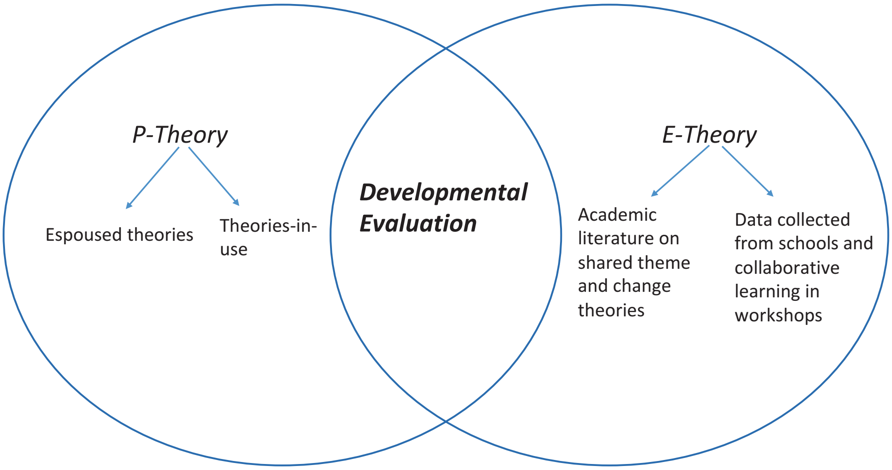

P-Theory
Espoused theories
Theories-in-use
Developmental Evaluation
E-Theory
Academic literature on shared theme and change theories
Data collected from schools and collaborative learning in workshops

Fig. 11.2 Evaluation through theory-engagement

through policy documents and what people say they do (and why), the latter can only be made visible through observations of behavior (i.e. teaching practices). Making these theories of action 'visible' is the main task of our review visits and workshops.

Schools found this element of the process in particular extremely useful. When evaluating the extent to which RiPR had influenced collaborative review practices 2 years after the pilot programme, one school said:

> *"we have found the concept of espoused theories and theories in action to be very powerful in a range of strategic reviews. We focused on this in subsequent reviews and the process of structured reflection here is what really moves practice forward."* (Highlands Primary School).

By linking evaluation theory to practice, we also add a further element to the process that will be described below. By recognizing that evaluators' evaluations in some ways compete with practitioner theories, it is through the reconciliation of these distinct theories that development comes about. Thus – and importantly – there is not an assumption that E-theory always trumps P-theory, rather that each may have their own merits and in combination have the potential to create new ideas and to refine ways of working.

6. Experimental research in the US has shown that collaborative school work when combined with external facilitation can lead to improvements in student achievement (Saunders et al. 2009). Evaluations of the RiPR clusters also tell us that participants value the role of the external facilitator, one stating that facilitation kept the review conversation grounded on evidence and theory. Skillful facilitation adds support, direction, increases research-engagement and provides critical friendship. This provides a challenge in other peer review models that are created by schools alone, since cluster working can have the negative effect of recycling

bad ideas or over reliance on anecdote rather than rigorous evidence. We know from the literature that not all school practitioners have the necessary evaluation skills or literacy and this diminishes the impact of internal evaluation (Nelson et al. 2015). Therefore, our process has always involved external (and we like to think, highly skilled) facilitation. Other research on collaborative school improvement particularly highlights the role for university facilitators in providing critical friendship to the network activities (Swaffield and MacBeath 2005). In RiPR, this means that we can remind participants to make judgments based only on the data gathered (and not their own favourite views) and to think back to what the academic research may add to the understanding of the situation. The impartiality of the university-based participant can sometimes make it easier for certain views to be voiced and to consider things from a completely different perspective.

7. As with other enquiry-based, evaluative approaches to school development, RiPR follows a cyclical approach to action and reflection. These four phases are outlined in Robinson's evaluation by theory engagement approach to leading change. First, school leaders agree on the problem to be solved, second, they make visible theories of action (espoused and theories in use), third, they compare the relative merits of current and alternative (new and desired) theories of action and, finally, they agree with teaching staff a new theory of action to be implemented (Robinson 2017). The acknowledgement by school leaders of teachers' theories of action – especially the reasons they give and values behind their actions, avoids leaders adopting a so-called 'bypass model' for leading change mentioned above (ibid, 2017). Having gone through one full cycle of RiPR, it is highly likely that participants will have a new and richer understanding of the issue they initially identified as 'the problem' that goes beyond merely refining strategies to the same issue. This redefinition of the problem to be solved is an example of what is called 'double loop' learning in the organisational learning literature (Argyris 1976).

The final stage of our last RiPR workshop concerns planning for implementation of the new approaches. Here, we emphasise the multi-faceted nature of the gap between current practices and desired future ones; this has been described as an 'implementation bridge' (see Hall 2013). Hall (and Hord's) work on implementing innovations in organisations suggests that there are behavioral, affective and adaptive components to the innovation use that go beyond practitioners merely 'using' or 'not using' the new desired strategies. Rather, the new ideas are often implemented mechanically and without understanding of the principles behind them; or some, we discover that staff are more 'ready' to implement change and recognize the rationale behind this than others. Finally, the ways in which the changes are enacted in practice fall along a sliding scale of expertise in terms of their enactment. These 'innovation configurations' (Hall 2013) can be co-constructed with leadership/teaching staff and used as rubrics to evaluate teacher practice in classroom observations. Schools found this element of the process more challenging, particularly in their initial encounters with it. However, a number drew on Hall and Hord's notion of Innovation Configurations (Hall 2013) to design observation tools to evaluate the extent to which new practices had been introduced.

8. Whenever schools are involved in collaborative work aimed at producing rich learning conversations, a culture of trust is essential (Stoll et al. 2006). Research has shown that high levels of trust aid complex information sharing and problem solving, shared decision making, and coordinated action (e.g. Bryk et al. 2010; Tschannen-Moran 2004). Brown et al. (2016) also show that high levels of trust are three times more important than other factors in fostering research engagement by teachers. Therefore, RiPR clusters come as self-formed; school leaders have generally worked together before and trust each other or have a good relationship that they wish to build upon and add structure to. Early in the process, we worked to co-construct explicit protocols with clusters so that they felt able to openly share their data and their weaknesses or issues that concerned them, rather than reacting defensively, hiding their shortcomings or seeking to demonstrate to their colleagues that they are competent and professional. This 'improving' rather than 'proving' focus provides a clear 'learning' rather than 'performance' orientation to our work Watkins (2010) and is vital to the success of the collaboration.

Below, we develop our rationale for how involvement in RiPR impacts on participants and their students. We do this by first by exploring in more depth the research base underpinning the key elements of evaluation theory and its relationship with evaluation policy. Secondly, we consider how RiPR supports the development of teacher collective efficacy.

11.5 Mechanisms for Impact of Research-Informed Peer Review

In this section we outline two key mediating processes that operate within our peer review model and show how these can lead to impact on school leader, teacher and pupil outcomes. The first, (using the example of RiPR collaborations that focus on feedback and assessment) occurs through an analysis of the school's assessment (evaluation) policies in relation to evaluation theory. For this we rely particularly on the work of Alkin and Christie (2004) on evaluation theory and, in particular, a more recent article by Christie and Lemire (2019) that makes the link to evaluation policy. We show with examples in RiPR work, how the explicit introduction of evaluation theory can have beneficial effects on school assessment policies. The second process we will look at suggests that the RiPR model may increase teacher collective efficacy, a factor found to have the highest level of impact on student attainment of any single strategy as measured in meta-analyses of educational interventions worldwide (e.g. Eells 2011). While we have not collected data on the latter, there is a clear rationale to suggest that RiPR processes and principles align with the aim of increasing teacher collective efficacy, and therefore future research looks very promising in this area.

11.6 Using Evaluation Theory to Inform Evaluation Policy

In Christie and Lemire's paper (2019) they argue for a purposeful integration of evaluation theory into organisational evaluation policies. In doing so, the authors argue that a number of potential benefits will accrue:

Their argument is as follows:

> *"..evaluation policies serve as a transformative mechanism for translating and perhaps more importantly situating evaluation theory in the organizational, political, and cultural context of an organization. In order to do this, in order to bridge the theory–practice divide, we argue that the theory integration has to be purposeful and explicit, emphasizing adaptation over adoption"* (Christie and Lemire 2019., p. 2).

Employing a similar rationale, we have encouraged an analytical, theory-informed approach to the evaluation of school's assessment policies. In RiPR workshop one, and then throughout the process, as mentioned in principle 5, above, we refer to Robinson's evaluation through theory engagement (Robinson 2017; Timperley and Robinson 2013). Robinson, in turn has been influenced by the groundbreaking work on organisation learning and change conducted in the 1970s by Argyris and Schön. In particular, we use the latter's simple conception of a theory of action (ToA) which states that any ToA is essentially composed of:

a) *the values and associated beliefs that explain*
b) *the observed actions and*
c) *the intended and unintended consequences of those actions"* (Argyris 1976).

We have found that by de-constructing school assessment policies into these three components, substantial learning takes place, often leading to decisions to transform school policies (and by corollary, practices). In the first workshop, we ask each school team to create a model to explain their school's assessment and feedback policies. As well as being an enjoyable team activity (we use lego, sweets, scissors, glue, marker pens etc. on a poster backing), we find that schools are able to use this modeling task to both clarify their thinking and to communicate their understandings and vision with others. Such sense-making processes (Stevens 2013) have been shown in the literature to help "managers [to] debate specific strategic challenges in a generative fashion" (Heracleous and Jacobs 2008, p. 310). Constructing mental models works by enabling those involved in the process to develop common meaning and purpose through a shared language or metaphor (Nonaka and Takeuchi 1995; Heracleous and Jacobs 2008). Importantly, this shared language is grounded in the props themselves, meaning that they can be used at a later date to re-create meaning for others. Therefore, we always photograph the models and encourage school participants to do the same. Many of our school teams have then repeated this process with teaching teams in their own schools (which we encourage).

Through explicit prompting, school teams visually represent their own school's assessment policies. We ask teams to explicitly show these as espoused theories of action (the talk), with beliefs, actions and consequences. As Christie and Lemire state: "Evaluation theory is who we are, what we preach, and (at least ideally) what we practice" (p. 2). While evaluation policies are situation specific, and reflect the particular nature of the context, personnel and particular operations of the organisation, with the primary purpose to guide actions and decision making, evaluation theories are context – free and serve as an aid to thinking about the purposes of the evaluation (Christie and Lemire 2019). A focus on evaluation theory can also reveal the philosophical orientation of the organisation as these authors elaborate through worked examples of three organisations' evaluation policies. The challenge for organisations then, once they have identified their evaluation theory, is to adapt and translate this into a policy that puts this into the school's specific context.

In RiPR workshops, school teams are asked to think about the people involved in assessment actions, the frequency of assessment and type of assessment, how they are conducted, by whom and why. They then reflect on what the intentions are (proposed desirable consequences) of the types of feedback given to students and the assessments of their learning. In the case of English schools, we have seen examples of how participants have realised that their policies contain a plethora of prescriptions to actions for teachers and a dearth of beliefs, values and educational justifications for these actions. The preferred consequences tend to be rather broad and unsophisticated assertions about maximising student learning). In one case in England, we have seen how this modelling activity led to a senior leadership team in a primary school to radically re-think and re-write their school assessment policies, co-constructing them in a learning exercise with their teaching team (see Godfrey and Brown 2019b). Their new policy was then trialled by teaching staff and evaluated during our peer review visit, late in the school year. This new policy made clear links to the literature on feedback and outlined the school's desire to encourage self-regulation and metacognition in their students, an aspect seen as highly effective in feedback practices (e.g. Hattie and Timperley 2007).

In the case of primary school teams in Chile, we saw a similar paucity in the school teams' understanding of teaching, learning and evaluation, particularly dominated by bureaucratic control and supervision and at the expense of a shared educational vision (see, Chap. 7). As with school assessment policy documents in England, these tended to emphasise prescribed ways and timings (and frequencies) for student assessment. More so than we saw in England, these assessment actions were tied to specific government legislation, giving more of an appearance of a policy dictated from the top than created at organisational level.

In the Chilean cases, this led to a transformation not only of an understanding of student learning, towards an idea of needing to allow more active involvement by students, but also in a parallel (and linked) need to develop an understanding of teachers as (active) adult learners. These changes came about as a result of an evaluative process that linked theory to the production and enactment of school evaluation policies.

11.7 Teacher Collective Efficacy

So far it has been difficult for the research to demonstrate a direct link between schools' involvement in peer review activities and improvement in student outcomes. Since most school peer review programmes work directly with staff at senior or (less often) middle leadership level, the effects of such involvement is likely to be indirect and mediated through a number of contextual and organisational factors. Nevertheless, we see a clear intention by those that create such programmes and for participants in them, to have an impact on student achievement. This relates to Gilbert's (2012) explanation of the moral accountability imperative of peer review working. One promising mechanism by which peer review involvement could work to achieve improved student learning outcomes is by increasing collective teacher efficacy (CTE).

Work in this area was pioneered in the 1970s by Albert Bandura. Bandura proposed that people's *beliefs* in their ability to achieve change through their own efforts and functioning (self-efficacy) had a powerful influence on their success in doing so. Furthermore, he saw that the same principle occurred at group level, coining the term 'collective efficacy'. Bandura defined this as:

> *"a group's shared belief in its conjoint capability to organize and execute the courses of action required to produce given levels of attainment"* (Bandura 1997, p. 477, from Donohoo et al. 2018, p. 40).

More recently, researchers have constructed scales to measure collective teacher efficacy (e.g. Goddard 2002). When school teams collectively feel confident that their own actions can, and do, have an impact on student learning (and attribute student learning less to factors like parental involvement, student SES or prior achievement of the school), then CTE is said to be high. Based particularly on doctoral work by Eells (2011), who conducted a meta-analysis of international studies on the impact of CTE, Hattie (2016) has now positioned CTE as the new number one factor that impacts on student achievement. With an effect size of 1.57, this is much higher than student prior achievement (0.65), socioeconomic status (0.52), home environment (0.52) or parental involvement (0.49) (Hattie 2016).

Strongly associated with high CTE is an environment where teachers share high expectations and a common language about student learning, rather than acting through compliance to teaching and learning regulations. Teachers see themselves as agents of change, as evaluators of their practices on student learning and as collaborators, working together to evaluate and improve students' learning. Thus, for teaching teams with high CTE, "success and failure in student learning is more about what they did or did not do, and they place value in solving problems of practice together" (Hattie and Zierer 2018, from Donohoo et al. 2018, p.41).

How can RiPR collaboration build collective teacher efficacy?

RiPR groups involve school senior leaders, other middle leaders and often teachers with no formal leadership experience. The workshops and peer review learning are followed by targeted opportunities for those teachers not initially involved to engage in similar learning back at school, followed by targeted implementation of new strategies in specific aspects of pedagogy. Further teacher collaboration is encouraged in the model, and made more likely since clusters are usually high in trust.

Thus, RiPR collaborations have the potential to develop collective efficacy in at least four connected ways:

1. By developing leaders' collective efficacy;
2. by helping leaders set a narrative focused on high expectations of (a shared idea of) student learning;
3. by collecting evidence of the impact of teachers' actions on student learning;
4. by increasing the density and quality of teacher collaborative networks.

1. Collaborative working among the leadership teams in peer review clusters can strengthen their own beliefs that they can positively affect student learning by leading teacher instructional learning. In the RiPR process, this occurs through the support, challenge and sharing of good practices during school visits, data collection and workshop discussions. Research, although still burgeoning, on collective school leader efficacy (CLE) has shown signs of a link between such leader CLE and leadership practices found to be effective. In one study, a supportive infrastructure for school leaders that helps to build school leader efficacy, is closely associated with principals' efforts at organizational redesign, especially building collaborative cultures and structures (Leithwood and Jantzi 2008, p. 521). The latter is supportive of the growth of CTE, particularly when school leaders feel confident in their ability to motivate teachers, to generate enthusiasm for a shared vision of the school, to manage school improvement, to create positive learning environment and to facilitate student learning and raised achievement on tests (Ibid, p. 512).
2. Where formal school leaders have considerable leverage is by controlling the narrative of the school. As Donohoo and colleagues state:

> *"If the narrative is about bus timetables, tweaks in the curriculum, and test schedules, this percolates through the school as the purpose of schooling—compliance to procedures. In such schools, students think learning is coming to school on time, sitting up straight, keeping quiet, and watching the teacher work. But if instead the narrative is about high expectations, growth in relation to inputs, what it means to be a "good learner" in various subjects, and what impact means, then teachers and students will think about learning in a different way. They will believe that learning is about challenge, about understanding and realizing high expectations, and that setbacks are an opportunity to learn"* (Donohoo et al. 2018, p.44).

We have seen in the case of Chilean adoption of this peer review model, that school leaders shifted from a compliance mode of leadership to one based much more on learning; and connecting teacher learning directly to student learning (see

Chap. 7 within). In the English experience, we have also noted that school leaders began to re-define their expectations of student learning and teacher feedback through re-modeled school assessment policies. For example, one school commented that:

> "*our self-review of feedback and marking led to a direct policy change and to a sequence of staff training sessions and focused observations. Our learning was recorded in a new policy document which we use continuously to inform the process of assessment, observation and planning. This has been a significant improvement in our practices.*" (Gearies Primary School).

3. Collecting evidence of the impact of new teaching strategies on student achievement can strengthen CTE (Donohoo et al. 2018, p. 44). RiPR school visits and the school's self-evaluations can provide ample data to this effect. This then reinforces future collective efforts and motivations to improve teaching strategies. This virtuous cycle is known as "reciprocal causality" (Bandura 1993, p. 44 in Donohoo et al. 2018). The approach of RiPR towards very targeted and precise data collection in classroom observations provides a particularly powerful tool in this respect since schools can then continue to use these methods in learning walks and peer observations or lesson study triads. School leaders have a role in creating "non-threatening, evidence-based instructional environments" (Donohoo et al. 2018, p. 42) that support this process at school level.
4. Collective teacher efficacy can be positively influenced through the improved networking opportunities for teachers involved in the RiPR process. This is supported through prior research using social network analysis. In a study of 53 elementary schools in a large Dutch school district, Moolenaar et al. (2012), found that the density of work-related and personal-advice networks affected teachers' perceptions of collective efficacy, which in turn was associated with increased student achievement (p. 258). Through RiPR collaboration, teachers directly involved in the process have the opportunity to network with colleagues in other schools. This research strongly suggests that school leaders should continue to encourage strong teacher relationships, as this would likely have a highly beneficial effect on CTE. The RiPR process strongly supports this model too, with both within school and between school teacher collaboration promoted during and after the cycle of reviews and workshops.

11.8 Conclusions and Discussion

The description of the RiPR model above explicates our key operating principles as a research-practice approach to joint practice development through a peer review process. Through successive iterations and adaptations, the model has been thoroughly concept-proven. Further thoughts on the mechanisms for impact have described two conceptual advances in our thinking: first, by linking evaluation theory to policy and secondly through increasing collective teacher efficacy.

Christie and Lemire's research suggests that the adaptation of evaluation theory to local school conditions is the key challenge to school teams in the improvement of their policies.

The implications of this are that the RiPR model should take an even more structured and focused approach to deconstructing school policies using more detailed descriptive and analytical categories (see p. 8–9, Christie and Lemire 2019). Undertaking this activity in workshop 2 could further enrich the mutual policy scrutiny exercise, leading to more specific suggestions for their refinement as well as improving participants' evaluation literacy.

The work on CTE provides food for thought for the RiPR model, since the potential gains for student achievement look considerable and because previous attempts to research peer review models have yet to make a clear link between these models and impact at student level (although one of the authors is currently involved in the largest trial of a peer review programme ever conducted[1]). In a Canadian context, Leithwood and Jantzi's research looked at how district leaders built a sense of collective efficacy among principals. In doing so they created a new four-item survey scale to measure leaders' collective efficacy beliefs about school improvement. These items, are listed below:

("To what extent do you agree that …")

1. School staffs in our district have the knowledge and skill they need to improve student learning?
2. In our district, continuous improvement is viewed by most staff as a necessary part of every job?
3. In our district, problems are viewed as issues to be solved, not as barriers to action?
4. Central district staff communicates a belief in the capacity of teachers to teach even the most difficult students? (Leithwood and Jantzi 2008, p. 512)

Districts encouraged higher school leader efficacy by emphasizing the priority they attached to student achievement, improving teaching and learning practices, and by providing a focus for school improvement efforts through cooperative working relationships with schools.

With RiPR collaborations occurring across different national contexts, these district (or equivalent) local contexts will be worthy of consideration in future analysis. The extent to which university facilitators can improve school leaders' collective efficacy – with or without district level alignment – will be interesting to observe. In the case of the experiences in England, we have seen that local authorities may no longer have the capacity to provide the type of support offered in the past and therefore newly formed alliances may be taking over this middle tier support level. These new partnerships, varying in formality, permanence and scale, provide a mediating variable that may have an indirect but potentially strong influence on the sustainability of collaborative improvements in participating schools and on how these filter down to teacher collective beliefs about efficacy.

---

[1] https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/schools-partnership-programme-spp/

A further caveat must be injected into the research on CTE; this is the way in which this variable is likely to be strongly mediated by other external factors. In particular, some research suggests a strong link between socioeconomically disadvantaged schools and lower levels of CTE. Belfi et al. (2015), citing a range of previous research, suggest several possible reasons for this, such as: less privileged student populations, lower student prior achievement levels, more student behavioral problems, lower levels of parent involvement, high student mobility rates, chronic student absenteeism and a poorer physical environment (p. 34). The suggestion is that these difficult circumstances have a negative effect on teachers' perceptions that can make a difference to their students' achievement. They cite further research to show that teachers in socioeconomically disadvantaged schools have lower educational expectations of their students than teachers in more socioeconomically privileged schools. We have seen some examples of this in our experiences in Chilean schools (see Chap. 12).

We can see a certain circularity in argument here, in that the very definition of CTE is to believe in one's own school team's ability to have an effect, *irrespective* of the contextual challenges. However, this appears much easier to achieve when such challenges are lesser in extent. Belfi and colleagues' research provides some optimism here however by positing a strong link between CTE and social capital theory (Bourdieu 1986), i.e., one that looks at the exchange of social resources (trust, support, norms and values) designed to leverage change. In particular, they looked at the relationship between CTE and school-based social capital (SBSC). They defined this as the social relationships that exist within schools and the resources that are exchanged through these relationships (Belfi et al. 2015, p. 35). Finding strong and significant statistical relationships between SBSC and CTE, they concluded that a focus on developing (trusting) relationships among teachers, students and parents, and by developing shared norms regarding good habits of schooling among teachers, students, and parents could be beneficial (Belfi et al. 2015, p. 42). While RiPR groups have not as yet involved parents to this degree, at least not explicitly, the involvement of students in developing a shared understanding of learning has certainly been evident in both English and Chilean cases. Further work on developing SBSC seems worthy of consideration to boost the chances of increasing CTE and ultimately student outcomes.

References

Alkin, M. C., & Christie, C. A. (2004). An evaluation theory tree. In *Evaluation roots: Tracing theorists' views and influences* (pp. 12–65). Thousand Oaks: Sage.

Argyris, C. (1976). Single-loop and double-loop models in research on decision making. *Administrative Science Quarterly, 21*(3), 363–375.

Baker-Doyle, K., & Yoon, S. A. (2010). Making expertise transparent: Using technology to strengthen social networks in teacher professional development. In A. Daly (Ed.), *Social network theory and educational change*. Cambridge, MA: Harvard Education Press.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117–148.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman and Company.

Belfi, B., Gielen, S., De Fraine, B., Verschueren, K., & Meredith, C. (2015). School-based social capital: The missing link between schools' socioeconomic composition and collective teacher efficacy. *Teaching and Teacher Education, 45*, 33–44.

Biesta, G. (2007). Why "what works" won't work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory, 57*(1), 1–22.

Black, P., & Wiliam, D. (2005). *Inside the black box: Raising standards through classroom assessment*. London: Granada Learning.

Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). New York: Greenwood.

Brown, C., Daly, A., & Liou, Y. H. (2016). Improving trust, improving schools: Findings from a social network analysis of 43 primary schools in England. *Journal of Professional Capital and Community, 1*(1), 69–91.

Bryk, A., Sebring, P., Allensworth, E., & Luppescu, S. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago: University of Chicago Press.

Christie, C. A., & Lemire, S. T. (2019). Why evaluation theory should be used to inform evaluation policy. *American Journal of Evaluation*. https://doi.org/10.1177/1098214018824045.

Donohoo, J., Hattie, J., & Eells, R. (2018). The power of collective efficacy. *Educational Leadership, 75*(6), 40–44.

Earley, P., & Porritt, V. (2014). Evaluating the impact of professional development: The need for a student-focused approach. *Professional Development in Education, 40*(1), 112–129.

Eells, R. J. (2011). *Meta-analysis of the relationship between collective teacher efficacy and student achievement*. Unpublished doctoral dissertation. Loyola University of Chicago.

Fielding, M., Bragg, S., Craig, J., Cunningham, I., Eraut, M., Gillinson, S., Horne, M., Robinson, C., & Thorp, J. (2005). *Factors influencing the transfer of good practice*. Nottingham: DfES research brief RR615.

Gilbert, C. (2012). *Towards a self-improving system: The role of school accountability*. Nottingham: National College of School Leadership.

Goddard, R. (2002). A theoretical and empirical analysis of the measurement of collective efficacy: The development of a short form. *Educational and Psychological Measurement, 62*(1), 97–110.

Godfrey, D. (2017). What is the proposed role of research evidence in England's 'self-improving'school system? *Oxford Review of Education, 43*, 433–446.

Godfrey, D., & Brown, C. (Eds.). (2019a). *An ecosystem for research-engaged schools: Reforming education through research*. Abingdon/New York: Routledge.

Godfrey, D., & Brown, C. (2019b). Chapter 6: Innovative models that bridge the research-practice divide. In D. Godfrey & C. Brown (Eds.), *An ecosystem for research-engaged schools: Reforming education through research*. London: Routledge.

Godfrey, D., & Handscomb, G. (2019). Chapter 1: Evidence use, research-engaged schools and the concept of an ecosystem. In D. Godfrey & C. Brown (Eds.), *An ecosystem for research-engaged schools: Reforming education through research*. London: Routledge.

Greany, T., & Higham, R. (2018). *Hierarchy, markets and networks: Analysing the 'self-improving school-led system' agenda in England and the implications for schools*. London: UCL Institute of Education Press.

Hall, G. E. (2013). Evaluating change processes: Assessing extent of implementation (constructs, methods and implications). *Journal of Educational Administration, 51*(3), 264–289.

Handscomb, G., & MacBeath, J. (2003). *The Research-engaged school* on behalf of FLARE, Essex County Council.

Hattie, J. (2016, July 11). *Third annual visible learning conference (subtitled Mindframes and Maximizers)*. Washington, DC.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.

Hattie, J., & Zierer, K. (2018). *Kenne deinen Einfluss!: "Visible Learning" für die Unterrichtspraxis*. Baltmannsweiler: Schneider Verlag Hohengehren GmbH.

Heracleous, L., & Jacobs, C. D. (2008). Understanding organizations through embodied metaphors. *Organization Studies, 29*(1), 45–78.

James, K. T., Mann, J., & Creasy, J. (2007). Leaders as lead learners: A case example of facilitating collaborative leadership learning for school leaders. *Management Learning, 38*(1), 79–94.

Leithwood, K., & Jantzi, D. (2008). Linking leadership to student learning: The contributions of leader efficacy. *Educational Administration Quarterly, 44*(4), 496–528.

Matthews, A. P., & Headon, M. (2015). *Multiple gains: An independent evaluation of challenge Partners' peer reviews of schools*. London: Institute of Education Press.

Moolenaar, N. M., Sleegers, P. J., & Daly, A. J. (2012). Teaming up: Linking collaboration networks, collective efficacy, and student achievement. *Teaching and Teacher Education, 28*(2), 251–262.

Muijs, D., & Rumyantseva, N. (2014). Coopetition in education: Collaborating in a competitive environment. *Journal of Educational Change, 15*(1), 1–18.

Nelson, R., Ehren, M., & Godfrey, D. (2015). *Literature review on internal evaluation*. London: Institute of Education.

Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. New York: Oxford University Press.

Robinson, V. (2017). *Reduce change to increase improvement*. Thousand Oaks: Corwin.

Robinson, V. M., & Timperley, H. (2013). School improvement through theory engagement. In S. Kushner & M. Lei (Eds.), *A developmental and negotiated approach to school self-evaluation* (pp. 163–177). Bingley: Emerald Group Publishing Limited.

Saunders, W. M., Goldenberg, C. N., & Gallimore, R. (2009). Increasing achievement by focusing grade-level teams on improving classroom learning: A prospective, quasi-experimental study of title I schools. *American Educational Research Journal, 46*(4), 1006–1033.

Sharp, C., Eames, A., Sanders, D., & Tomlinson, K. (2005). *Postcards from research-engaged schools*. Slough: National Foundation for Educational Research.

Spillane, J., Healey, K., & Kim, C. (2010). Leading and managing instruction: Formal and informal aspects of elementary school organization. In A. Daly (Ed.), *Social network theory and educational change*. Cambridge, MA: Harvard Education Press.

Stevens, J. (2013). Design as communication in microstrategy: Strategic sensemaking and sensegiving mediated through designed artifacts. *AI EDAM, 27*(2), 133–142.

Stoll, L., Bolam, R., McMahon, A., Wallace, M., & Thomas, S. (2006). Professional learning communities: A review of the literature. *Journal of Educational Change, 7*(4), 221–258.

Swaffield, S., & MacBeath, J. (2005). School self-evaluation and the role of a critical friend. *Cambridge Journal of Education, 35*(2), 239–252.

Tschannen-Moran, M. (2004). *Trust matters: Leadership for successful schools*. San Francisco: Jossey-Bass.

Watkins, C. (2010). *Learning, performance and improvement* (Research Matters (34)). London: IOE.

Wilkins, R. (2011). *Research engagement for school development*. London: Institute of Education.

Dr. David Godfrey is an Associate Professor in Education, Leadership and Management at UCL Institute of Education in London and the programme leader for the MA Educational Leadership. He was co-director of the Centre for Educational Evaluation and Accountability until 2018 and was a lead inspector for the Independent Schools Inspectorate. An advocate of research-informed practice in education, his projects and publications include research-engaged schools, school peer review, inspection systems and lesson study. In July 2017, David was acknowledged in the Oxford Review of Education as one of the best new educational researchers in the UK.

Karen Spence-Thomas is a Programme Leader in the London Centre for Leadership in Learning at UCL Institute of Education. She has held teaching and leadership roles in London schools and now specialises in designing and facilitating tailored professional development programmes in the UK and internationally. She has a particular interest in creating capacity for R&D in schools and co-leads the IOE R&D Network. Recent and current projects include: the EEF-funded Research Learning Communities (with Chris Brown); ESRC funded Middle Leaders as Catalysts for Change (with Louise Stoll); NCTL Research Themes Project (which she co-led with Carol Taylor). Recent publications include: 'Understanding Impact and the Cycle of Enquiry' with Carol Taylor in C. Brown (ed, 2015) *Leading the Use of Research and Evidence in Schools.*

Chapter 12
Changing School Leaders' Conversations about Teaching and Learning through a Peer Review Process Implemented in Nine Public Schools in Chile

Mónica Cortez, Fabián Campos, Carmen Montecinos, Jorge Rojas, Marcela Peña, Jorge Gajardo, Jorge Ulloa, and Charles Albornoz

Abstract This chapter describes the design, implementation and outcomes of a professional development model for school leaders based on an adaptation of the Research-informed Peer Review (RiPR). Leadership teams from nine schools participated in the School Inquiring and Learning with Peers (SILP) program which focused on strengthening school leaders' capacities to analyze and reflect on the feedback classroom teachers provided to students. The model includes learning settings for network level and school level activities to support the transfer of learning achieved at the network level into leadership practices. The mediation of the university team, the participation of peers and the availability a broad range of evidence generated deeper conversations about the complexities of problems of practice to be worked on. Findings showed that participants developed a shared language of learning, made learning visible in their classroom observations, shifted their understandings of how to support teacher's professional development, and how to use evidence to inform their decision-making.

This work has been possible thanks to the collaboration of the Leadership Center for School Improvement, LIDERES EDUCATIVOS, and David Godfrey of UCL Institute of Education. The authors are grateful for the partial funding provided by CONICYT FB003.

M. Cortez (✉) · F. Campos · C. Montecinos
PCentro LIDERES EDUCATIVOS PUCV, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile
e-mail: monica.cortez@pucv.cl

J. Rojas · J. Gajardo · J. Ulloa
Facultad de Educación, Universidad de Concepción, Concepción, Chile

M. Peña · C. Albornoz
Saberes Docentes, Universidad de Chile, Santiago, Chile

© Springer Nature Switzerland AG 2020

D. Godfrey (ed.), *School Peer Review for Educational Improvement and Accountability*, Accountability and Educational Improvement,
https://doi.org/10.1007/978-3-030-48130-8_12

12.1 Introduction

A key task for school leaders is the creation of conditions that foster continuous improvement of organizational processes and results. Crucial among these is the implementation of opportunities for teacher learning and development, as this has been shown to be an instructional leadership practice that has the greatest impact on student learning outcomes (Bolivar 2015; Robinson et al. 2009). It is highly relevant, therefore, to develop powerful professional development opportunities for school leaders, carefully designed for the acquisition of specific skills and knowledge associated with this practice.

Sociocultural theory of learning proposes that the knowledge that sustains the work of professionals has an active, situated and distributed character (Lave 1988). This knowledge is constructed through social interactions in the context in which it is applied and generated in both the academic world and in the workplace (Borko 2004). A process of professional learning located in the performance of authentic leadership tasks involves the construction or reconstruction of: (a) existing patterns and rules regarding “what”, “why” and “how” the leadership work is carried out; (b) knowledge and skills to address typical problems or decisions when solving prototypical and relevant problems of practice; and (c) knowledge that simultaneously changes leadership practices and the context in which these are deployed. That is to say, learning implies the generation of new social and material configurations in the workplace (Montecinos and Cortez 2015).

Research on effective professional development (PD) is robust in identifying features that promote changes in practitioners’ knowledge, skills, dispositions, and practices (Kalinowski et al. 2019). Among these, the duration of the program such that it is a process and not an event, coherence with other activities designed to enhance practitioners’ performance, a link to participants’ own experience, interests and needs, and the involvement of external experts to expand conceptual understanding. Considering that teachers teach as they are taught, if PD seeks to mobilize teachers to implement deep learning, the PD program needs to foster active learning, collaboration, opportunities to practice with feedback and reflection on this practice and feedback.

These characteristics are present in professional development programs based on the implementation of a cycle of inquiry, action and reflection, supported by a community of peers. Learning through networks of schools allows articulating professionals from different institutions to advance together in the improvement of the learning outcomes achieved by students (Chapman et al. 2008). In fact, the evidence suggests school leadership teams have a greater impact on student learning outcomes when they learn with other leadership teams committed to an improvement inquiry cycle (Chapman and Muijs 2014).

This chapter describes the design, implementation and outcomes of a professional development model for school leaders that ascribes to these principles of professional learning. School Inquiring and Learning with Peers (SILP)[1] is based on an adaptation of the Research-informed Peer Review (RiPR) model developed by Godfrey (see Chap. 11 in this volume by Godfrey and Spence-Thomas; Godfrey and Brown,2019).[2] This model involves gathering evidence for school self-evaluation in collaboration with other school leaders, with the support of university-based facilitators who bring external expertise. Peer review is expected to deepen knowledge around the topic of inquiry, leading to new practices in the participating schools as participants revisit their theories of action in light of the evidence that is collected and analyzed through the inquiry cycle. Leadership teams from nine public schools (N = 27) participated in the implementation of SILP which focused on strengthening school leaders' capacities to analyze and reflect on the feedback classroom teachers provided to students, in order to enhance teachers' effective use of feedback (Cortez et al. 2019).

The chapter is structured into four sections. First, considering that school leadership is sensitive to cultural and policy contexts (OCDE 2017), we briefly highlight recent policies that orient the work priorities of Chilean principals. Next, we describe the key components of SILP, highlighting the learning settings[3] designed to mobilize participants' critical and reflective thinking about teaching and learning and their leadership practices that support or hinder these processes. The third section reports topics of conversation that emerged as participants interpret the evidence gathered during the peer review process, their own learning and how they used this learning to strengthen the use of teacher feedback at their schools. Lastly, we discuss the value and challenges of the different learning settings within SILP to promote collaborative learning among public schools in Chile.

12.2 Public Schools Leaders' Work in Context

Following an outline of the structure for the provision of educational services, we offer an overview of recent policies that have: (a) promoted the participation of private providers associated with high levels of segregation and inequality, aspects that characterize Chilean education; and (b) increased accountability to orient the work of public school practitioners, as more decision-making over pedagogical matters has been transferred to the school level.

---

[1] In Spanish the peer review model is called Centros Escolares que Indagan y Aprenden Juntos (CIAJ).

[2] See also Chap. 11 within this volume.

[3] A learning setting refers to the actions, material and conceptual tools and infrastructure through which learning is enacted, new practices emerge and how the interactions among individuals meaning making potential is mediated (Säljö 2009).

12.2.1 Structure of the System

The coordination and regulation of the education system from kindergarten to tertiary education is the responsibility of the Ministry of Education. Currently, students complete 12 years of compulsory schooling structured into 8 years of primary education and 4 years of secondary education (covering about 90% of the school-age population from 6 to 18 years old). Secondary education includes 2 years of a common curriculum and 2 years of a differentiated program: technical-vocational and scientific-humanist. In response to primary teachers' insufficient disciplinary preparation, this structure was supposed to change by March 2018 to include six primary and six secondary grade levels. This change, however, has not been fully implemented and most elementary schools serve students grades Kindergarten through 8. Universities offer initial teacher preparation, and primary teachers are generalists in the various subject areas covered in the national curriculum; secondary teachers are subject specialists.

12.2.2 Market Model for the Provision of Education

Chile is credited as the first country to install a market model for the provision of educational services in the early 1980s. Parents' right to choose is guaranteed in the constitution, opting among four types of schools: (a) municipal (public), financed through a state voucher; (b) subsidized-private, financed through the same state voucher; (c) municipal, administered by a private provider financed through same state voucher; and (d) private, fully financed by families. The state financial contribution via de voucher is based on the school's average daily attendance and school enrolment. This is an important contextual challenge in Chile as public schools within the same municipality compete with each other to increase enrolment.

After over 30 plus years of a quasi-market model, there has been a steady expansion of private providers and by 2017 this sector accounted for close to 64% of the enrolment in primary and secondary grades. Privatization has contributed to a highly segregated educational system with unequal learning opportunities (Carrasco et al. 2017; Valenzuela and Montecinos 2017). This has been created through the use of competition among schools as an incentive to improve quality, parental choice, student selection processes and family co-payment in the private sector schools (as many as 80% of all private subsidized schools were charging families a monthly tuition by 2013). Chile, according to the OCDE (2017), remains one of the most unequal member countries; this inequality is reproduced by the educational system as students from different socio economic groups attend different schools, which in turn achieve different outcomes (Valenzuela et al. 2014). Public schools concentrate a large proportion of low-income students and these schools also tend to be among those repeatedly attaining poor results on the national testing system of school quality (SIMCE). Notwithstanding, among Latin American countries participating in Pisa 2015, Chile attained the best performance (OCDE 2017).

12.2.3 Increased Decentralization and Accountability

Together with measures that sought to foster greater decentralization of pedagogical decisions a number of recent policies have increased external accountability. The *Preferential School Subsidy Law* (SEP) passed in 2008 provided additional funding, above the regular voucher, per low-income student enrolled at the school. This additional funding per low-income student enrolled was later increased through the *Law for School Inclusion a*pproved in 2015. SEP funding is tied to a school improvement plan stipulating targets to be met each year and the municipal department is responsible for their correct implementation. This affords schools greater autonomy to chart improvement paths, though they also became subject to greater levels of external accountability as they were required to account for the progress attained as well as the money received.

With the introduction of the *National System for Quality Assurance of Education* in 2011, the Agency for Quality Education became responsible for the development and administration of the national testing program (SIMCE) through which students' academic achievements are measured, including other non-cognitive measures of students' development. Largely based on SIMCE results, each school is categorized into one of four levels: high, middle, middle low, and insufficient. After 4 consecutive years in the insufficient category, representing 8% of all schools by 2018 and 17% of schools serving low-income students,[4] schools receiving the state subsidy should be closed by the end of the fifth year. Additionally, the Agency conducts school inspections targeting schools with poor results in SIMCE, orienting their improvement through a set of performance indicators. The Ministry of Education is charged with providing these schools direct, or through a consultancy service, technical support. Schools that achieve good results are also visited to provide authentic examples of how effective schools are implementing these indicators. By law, this agency is required to make these results publicly available to inform parents' choice.

The Education Superintendence was also created by as part of the quality assurance system. It is responsible for inspecting schools to safeguard adequate use of public funding as well as to ensure compliance with all legal requirements. This institution is mandated to receive and resolve complaints from parents (educational consumers) who claim their educational rights have been trespassed by the school's actions or omissions.

---

[4] Source: Agency for Quality Education. Retrieved from http://archivos.Agenciaeducacion.cl/Policy_brief_CD_2018.pdf

Law of Educational Quality and Equity (LQE), passed in 2011, provided new regulations to strengthen school leadership by, among other things, affording them greater autonomy over staffing and budget matters. Principals are required to sign a 5-year contract with their municipal department of education with predefined results to be achieved each year. This law orients principals to focus on instructional leadership responsibilities, in addition to administrative tasks, through the implementation of the school improvement plan required by the *Preferential School Subsidy Law* (Montecinos et al. 2015).

The National System for Teachers' Professional Development Law, approved in 2016, enhanced teachers' working conditions. This law provides salary raises associated with teachers' performance on examinations of their knowledge, a decrease in the number teaching hours, allowing more time for lesson preparation and collaboration with peers, a universal induction program for newly qualified teachers as well as a new professional development model. Whereas until 2017 professional development was largely centrally defined by the Ministry of Education, the new model is school-based, charging the school principal with the tasks of developing a professional development plan that is responsive to the school's challenges.

Taken together, these reforms entail strengthening leadership capacity at the school and at the intermediate level to enhance systemic coherence. Research evidence indicates that schools by themselves cannot address the challenges posed by the most vulnerable and complex contexts and competition is not a leverage for improvement in cases like Chile, where the system is highly inequitable and segregated by social class (Hattie 2008; Ahumada-Figueroa et al. 2016). Network has been suggested as a more effective route and SILP uses this principle to leverage school leaders' capacities to respond to the requirements set by the National System for Teachers' Professional Development, SEP and LQE laws.

12.3 Schools Inquiring and Learning with Peers Model

The model was implemented with leadership teams from nine public (municipal) schools located in three regions of Chile (Valparaíso, Biobío and Metropolitana), were configured into a network of three schools per region. Table 12.1 presents key characteristics of each school as well as the data sources during the peer review process. Within each region, all schools were part of the same municipal government and within each network schools in different Quality Agency categories were included. The municipal director of education and all schools selected initially by this director were invited to participate on a voluntary basis. All professionals within each school also volunteered to participate and signed a consent form agreeing to the participation in research activities and data collection.

The school level team comprised the principal, the curriculum coordinator, and a teacher leader selected by the principal (N = 27). Among them, 93% are female and 59% had been at their current position for less than 3 years. In each region, a university-based team affiliated with the Leadership Center for School Improvement

Table 12.1 Characteristics of Participating Schools and Data Scources

							Data sources during peer review		
School	Region	Grade levels	Number of teaching staff[a]	IVE[b]	2018 enrolment[c]	Quality agency category[d]	Number student interviewed	Number of teachers interviewed	Number of classroom observed
S1	Valparaíso	K-12th	61	K-8th: 90%	654	Middle low	16	9	12
				9-12th: 89%					
S2	Valparaíso	K-12th	46	K-8th: 89%	533	Middle	12	9	16
				9-12th: 85%					
S3	Valparaíso	K-12th	44	K-8th: 91%	471	Middle low	15	9	15
				9-12th: 100%					
S4	Metropolitan	K-8th	55	89%	1082	Middle	6	6	14
S5	Metropolitan	K-8th	29	88%	482	Middle	6	6	18
S6	Metropolitan	K-8th	62	94%	929	Middle	6	6	19
S7	Bíobío	K-8th	22	100%	77	Insufficient	9	9	9
S8	Bíobío	K-8th	22	94%	98	Insufficient	11	7	7
S9	Bíobío	K-8th	26	96%	166	Middle	9	10	9

[a] Source: http://www.mime.mineduc.cl/

[b] IVE indicates the proportion of low income students students. Source: https://www.junaeb.cl/ive

[c] Source: http://datos.mineduc.cl/dashboards/19738/bases-de-datos-resumen-de-matricula-por-ue/

[d] Source: https://www.agenciaeducacion.cl/orientacion/categoria-desempeno/

facilitated the workshops, participated in the school visits and led follow-up sessions in each school. This team's key contributions were to explain and facilitate the process, provide research-based information regarding effective use of feedback, act as a critical friend problematising participants' theories of actions and design learning activities to foster inquiry and collaborative reflection.

12.4 Components of the Schools Inquiring and Learning with Peers Model

As can be observed in Fig. 12.1, the components largely replicate the process flowchart described by Godfrey and Brown (2019) for RiPR, and as described also in Chap. 11 in this volume. The model has four broad components: alliance with a municipal department of education, workshops at the university, school-to-school visits, and within school workshops. We highlight two key differences between the model developed by Godfrey and SILP. First, we established an alliance with a municipal department of education to ensure their support for the participation of school teams. The university team also engaged in additional in-school workshops with each leadership team and teachers in order to scaffold the transfer of learning into practice.

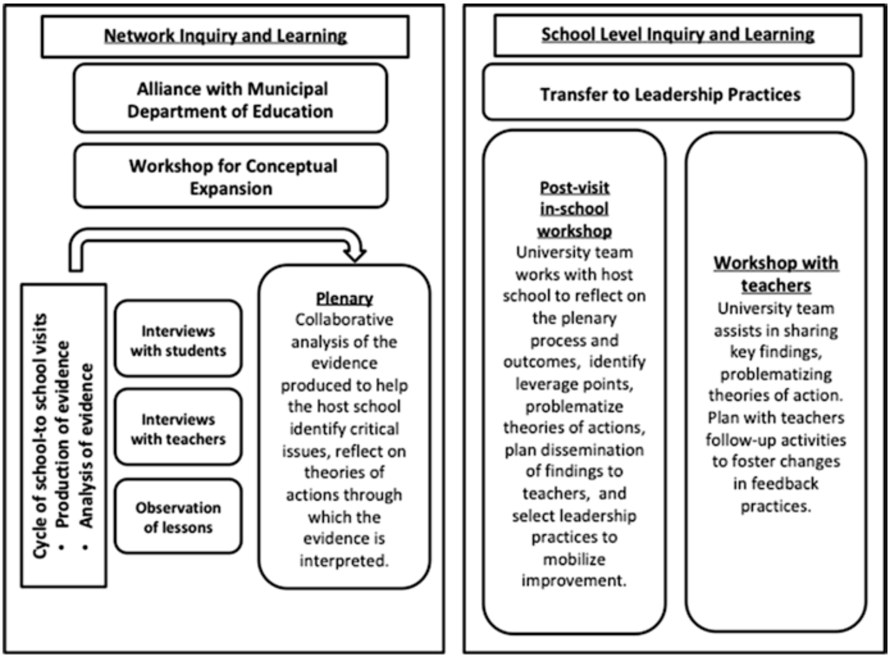

Network Inquiry and Learning
Alliance with Municipal Department of Education
Workshop for Conceptual Expansion
Cycle of school-to school visits
• Production of evidence
• Analysis of evidence
Interviews with students
Interviews with teachers
Observation of lessons
Plenary
Collaborative analysis of the evidence produced to help the host school identify critical issues, reflect on theories of actions through which the evidence is interpreted.
School Level Inquiry and Learning
Transfer to Leadership Practices
Post-visit in-school workshop
University team works with host school to reflect on the plenary process and outcomes, identify leverage points, problematize theories of actions, plan dissemination of findings to teachers, and select leadership practices to mobilize improvement.
Workshop with teachers
University team assists in sharing key findings, problematizing theories of action. Plan with teachers follow-up activities to foster changes in feedback practices.

Fig. 12.1 Components of the schools inquiring and learning with peers model

These components included several learning settings designed in pursuit of the following objectives:

1. *Systemic improvement*. In each region, the local university team developed a working alliance with a municipality aiming to support systemic improvement through networked learning. The municipality's director of education and the deputy for technical-pedagogical coordination first analyzed the program to determine its fit with their improvement plan and identified potential schools. Later, they organized a meeting with these schools' principals and the university team to explain the program components, objectives and the requirements of participation. During the course of the implementation of the SILP, the deputy coordinator attended the workshops and periodic meetings were held with municipal-level educational leaders to account for the work conducted and the challenges to be addressed at the municipal level. Additionally, their involvement sought to ensure that these teams were not called to participate in other initiatives that would add too many interventions within the school improvement plan or preclude them from attending scheduled SILP sessions.
2. *Conceptual expansion.* This involves learning to interpret evidence produce through the peer review process using new concepts, refining understandings of concepts already in-use by paying attention or becoming aware of features previously ignored, or restructuring by establishing new relationships among concepts. For example, when participants recognized that classroom observation should not only make teaching visible but also pupil's learning visible, we noticed there was conceptual expansion. The cycle included conducting two workshops prior to the schools visits that sought to expand and enrich participants' knowledge of the use of effective feedback in the classroom as well as the use of systematic production of evidence of teachers' practices and students' learning. In addition, these workshops delved into the theories of action and addressed sociocultural perspectives on teacher learning and development. Conceptual expansion is infused through all the activities as new conceptual understandings are used to interpret evidence as well as a lense to observe lessons.
3. *Strengthening analytical and reflective abilities to interpret evidence and make evidenced-informed decisions*. School to school visits lasted about 7 h each and followed a protocol through which peers and the host school team produced evidence. The host school had total control over which lessons in different grade levels and different subjects areas were observed and students and teachers to be interviewed. Observations followed a structured protocol that focused on the learning goals of the lesson as well as on the different types of feedback provided (Hattie and Timperley 2007). Interviews focused on asking about conceptions of learning (i.e., what do good students do, what makes a good learner), conceptions of teaching (i.e. characteristics of effective teachers) and conceptions of feedback (i.e., what is feedback, what types of feedback is provided by your teachers).

In a plenary session, lasting between 90 and 120 minutes, this evidence was shared to describe what was observed and what students and teachers said. This plenary session turned out to be a powerful instance for making visible to the host school the areas in need of improvement. However, the amount of information that needed to be processed and detailed attention to the unique school context (culture and structure) made this collective instance insufficient to unpack personal theories of action and identify concrete school-level leadership practices to mobilize improvement. To address these limitations, a post-visit in-school workshop was introduced to the SILP model. The university team facilitated this workshop seeking to problematize theories of actions underpinning how the evidence was interpreted. This instance allowed us to appreciate the difficulties that senior school leaders encountered as they thought about how to transfer into practice what they had learned in the cycles of inquiry.

4. *Modeling and support to enlist teachers' commitment to improving their feedback practices*. The difficulty presented by the school teams when thinking about how to get teacher buy-in, led the university team to offer to work with them to plan and deliver the workshop with teachers. This workshop focused on: sharing the findings from the school visit, problematizing beliefs about effective feedback and learning, conceptual expansion, and collegial conversations to develop a shared vision of effective learning and teaching at the school.

12.5 Key Findings

Data sources include the transcripts of the conversations that took place during the plenary sessions and in-school workshops with the university team. These were read multiple times by all researchers, coding them through a content analysis process (Stake 2010). Three sessions brought together the three university teams to collectively analyze the results of the coding. Findings are structured to reflect changes in participants' understandings of the following issues:

How is learning understood and enacted when teachers work with pupils and when interpreting the evidence produced at the school?

How is teacher professional development understood and enacted as school teams interpreted the evidence and planned for transfer of learning into practice?

As a result of their participation in SILP, what skills do these school leaders claim they have developed?

How do participants evaluate the various components in SILP's design and implementation?

12.5.1 Vision and Approach to Student Learning

Two changes were observed over time: (a) from the absence of shared language to talk about learning toward a shared vision of the learning approach to be promoted at the school and (b) from a focus on teachers and teaching to the recognition of students as protagonists. These two changes are part of the mind frames of effective teachers defined by Hattie (2012): develop a common and shared language of learning which is understood by all and talk more about how pupils learn than about teaching.

Developing a shared language and vision about learning. This finding reflects the goal of conceptual expansion that was afforded through the workshops and plenary discussions of the evidenced produced through the school visits. Initially, concepts such as learning, assessment and feedback were equivocated and participants recognized their meanings were taken for granted. As noted in the next interview excerpt, senior leaders acknowledged that in their instructional work with teachers, explicit conversations about learning were missing:

> I observed that when we were discussing the evidence, different people understood learning differently, some were talking from personal experience and others were describing the practices observed, whereas others were echoing theory. Therefore, I think conversations about learning, what are we going to understand and how are we going to describe learning processes, have not occurred (Curriculum Coordinator, S8, Bíobío Region, School Visit 2).

When talking about these concepts some participants used a normative vision pointing out the instruments and protocols they used to monitor and regulate teaching. They became aware that the use of these tools did not consider sense making and the moral purpose of the work of educators. Rather, the focus was on satisfying demands codified in several performance indicators through the various external accountability demands. In the following excerpt, a participant notes that when discussing underachieving students, the problem to be solved was to figure out how to promote the student from one grade level to the next as the Ministry directive is that students should not repeat a grade level. In doing so, attention shifted away from understanding why the particular students is not achieving as expected:

> At the end of the day we are focused on results, on what the standard is telling us, that is our focus. We are not attending to students' progress, to whether they are learning. We work toward the test and to meet performance indicators through which the Ministry assesses the school. Yes, to external accountability (Principal, S3, Region Valparaíso, Visit 1).

In the absence of substantive conversations, a shared vision regarding learning, assessment and evaluation was missing. As exemplified in the next excerpt, the review process helped schools recognize the need to develop a professional culture focusing on improving student outcomes rather than on meeting external requirements.

> We need to have a shared vision of learning to understand the pedagogical interactions and the feedback required. We are missing the fundamental elements; what is understood by learning, what do we want to achieve and, what do we want our students to learn? (Principal, S3, Region Valparaiso, Workshop 4).

From a focus on teachers and teaching to the recognition of students as protagonists in learning. In the first workshop participants were asked to map how assessment and feedback was used in their school. It became evident that they were exclusively focused on teaching. When conducting classroom observations, for example, their protocol looked at what teachers did without simultaneously attending to what students did during these lessons.

> Today, and this week, there has been so much learning. I had never realized, for example, that when I did classroom observations I never paid attention to feedback. Today I have a different perspective on classroom observation; we need to work on effective feedback, on assessment, on students' and teachers' conceptions of learning. Principal, S9, Bíobío Region, Visit 3).

The review processes put in sharp focus that in their schools, instruction was teacher-centered and that a passive view of the learner was promoted. This was disheartening because several actions had been taken to foster constructivist-learning environments as the national curriculum follows this approach. During the school visits, students were interviewed to understand their perspectives of instruction and learning and through classroom observation evidence of the use of feedback was obtained. What was often the case was that lessons involved students in activities that did not afford opportunities for quality feedback.

In their conversations with students, participants were surprised to learn that for a large majority of the students learning was equated with obtaining good grades. Additionally, for students a characteristic of a good learner is to behave well and not cause disruptions. The following excerpt of a conversation during the plenary following the first school visit is illustrative:

University team member: When you interviewed students, what were the skills they identified as key to being a good learner, a good student?

P1: *get good grades*
P2: *to not use swear words, not fight with classmates*
P3: *nothing related to academic* [abilities]
(Plenary, Region Valparaíso, Visit 1)

These students' understanding of learning, seemed well aligned with teachers' concerns and professional conversations:

> Our classes are flat, we worry too much about behavioral management- In our faculty meetings we spend a lot of time talking about students' disruptive behaviors (Principal, S3, Valparaíso Region, Visit 1).

During the plenary sessions, participants began to develop a vision of learning that called into question their teaching practices. Particularly the teachers in the teams identified a need to change their role in the lessons:

> Now, I think that we have to realize that we are mediators, that is a change, to be a guide because with the methodology we are now using we are the protagonist, and the student just listens (Teacher, S4, Metropolitan Region, Visit 3).

> The kids want to participate and we sometimes leave them on the other side of the street. We are the leading actor but when learning they should have a much more protagonist role (Teacher, S2, Valparaíso Region, Visit 1).

12.5.2 Understandings of Teacher Learning and Development

In the conversations documented overtime in the plenary sessions and in workshops, we observed a shift in how teacher learning was understood. Initially it was approached from the logic of training by external experts and more supervision. As the process moved along, we observed school leaders question their understandings of how their practices helped teachers grow and develop as professionals. In their new understandings they began to articulate the features of effective professional development.

After the first school visit, during the plenary as well as during the in –school sessions, participants identified teachers as the problem to be solved. Teachers were characterized as outdated, lacking professional knowledge and in some cases lacking interest or commitment to improve. With this quick diagnostic, the school teams right away wanted to take remedial actions such as hiring a consultant to increase teachers' capacities. The consultant would prescribe what good teaching entails and teachers would the do as they were told. Additionally, to improve the quality of teaching, the senior leadership team needed to increase surveillance and control:

> [When I arrived to this school] we thought that pedagogical guidelines were already established thus we did not want to introduce changes, we thought teachers had adequate knowledge about teaching methods and assessment. However, the interviews conducted [by peers] have been very positive because teachers could talk more freely as compared to when they talk with the principal. We became aware of the need to develop an evaluative culture with clear guidelines and we must work to institutionalize certain practices (Principal, S8, Bíobío Region, Visit 2).

When this theory of action is problematized during the first workshop, participants begin to acknowledge that supervision has not generated professional learning among teachers:

University team: *What evidence do you have that by having the curriculum coordinator revise all of the tests developed by teachers to assess students, these tests have become better over time?*

Principal 2: *better?*

University team: *If a year from now I come back to your school, will I find that that teachers' test still evidence the same problems or will those problems be resolved?*

Principal 2: *No, they repeat, they repeat what they are doing (…) so* [I] *must constantly revise* (Principal, S3, Valparaíso Region, Workshop 1).

School teams began to problematize their explanations for the prevalence of unchallenging lessons delivered to students, where rote repetition was the predominant way of knowing if students had learned. The problem was seen as entailing greater complexity, involving the creation of opportunities for teachers to critically reflect on their theories of action as a starting point to create a need for professional development. This shows an initial understanding that teacher learning is active, social and situated in the workplace, similar to their own learning experiences with SILP:

> We the leadership teams are very clear in terms of what learning should be pursued by the school, but we have not been able to get teachers to speak with this same intensity. I think it has to do with the fact that we have been part of a process of experimentation [referring to SILP] (…). If we do not want teacher robots, student robots we need a protected process where we are allowed to make mistakes (Principal, S1, Valparaíso Region, Visit 3).

> When you observe two teachers in the same subject area, in similar grades teaching so differently, you have to question that teachers work alone in the classroom. Teacher training is not always the answer, but we do not have time to go and observe how a peer teaches. I think there is a lot of richness in different practices and if teachers could see them they could reach their own conclusions about what needs to change (Principal, S5, Metropolitan Region, Visit 2).

During the first plenary session a principal said that she had previously heard that giving and receiving feedback among peers was very effective. She had her doubts, as one's ego would get involved, making one less receptive to peer criticism and finding little truth in the feedback provided. In the last workshop, she mentioned that she had changed her understanding after implementing peer observation among teachers in her school:

> Teachers broached the topic and we had to work hard to make it work, we worked with them, more than observation and evaluation, it was a process of feedback among peers. I think it was very effective and it was a need that emerged from teachers (Principals, Metropolitan Region, School 4, Visit 3).

12.6 Professional Skills Development Reported by School Teams

We analyzed what participants said regarding what they were learning throughout the implementation of the various components of SILP. Two main patterns emerged. First, they became aware of how to conduct more rigorous observations and how to use evidence to inform their decision-making. Second, the availability of evidence allowed them to engage in deeper conversations with peers, looking at problems of practice with greater complexity and engaging in critical reflection.

Observation and the use of evidence. The use of observation protocols that had a very precise focus and provided a tool to apply the concepts addressed in the workshops strengthened their ability to observe. By focusing on the quality of feedback they had specific evidence of an aspect of classroom instruction that could assist in defining a focus for professional development. The introduction of a classroom observation tool that directed the observer's attention to learning was recognized as an important enhancement to what they had been previously doing. This focus allowed them to better understand the distance between the learning goals that teachers had defined and the opportunities they afforded to students to reach those goals:

> I have learned to change the focus of my observation. This has been very important because I have been doing classroom observations but had not made visible how teachers use feedback and how they orient students' learning (Principal, Bíobío Region, School 9, Visit 1).

> When we talk, we know what we need to say because the context is telling us what we should say and think about these issues [teaching and learning]. But in practice it does not happened, we do not see kids working in groups, we are not teaching kids to work with each other, to learn with others, to discuss with others (Principal, Metropolitan Region, School 6, Visit 3).

By attending to students' engagement, participants also identified issues of inclusion and equity. Two key topics emerged from their conversation: whether boys and girls were having the same opportunities and whether teachers were making sure that all students attained the learning goals. During the plenary a participant commented that only boys raised their hand when a teacher asked a question, that the teacher did not encourage girls to answer and the same three or four boys were the ones the teacher called on over and over again. In another school, the peer had interviewed a ninth grade girl who told her she had not yet learned how to divide. She noted that the girl said this expressing frustration; she added, "*I think this girl cares about learning; I became frustrated and I left feeling a pain in my gut.*" (Plenary, Valparaíso Region, Visit 2).

Strengthening their analytical and reflective skills when using evidence. The availability of evidence produced through multiple sources by multiple people across school grades and subject areas created an opportunity to engage in meaningful professional conversation about teaching and learning. In these discussions participants were opened to revisit their theories of action. The most powerful evidence to break strongly held beliefs came from students' voices. Looking at the classroom from the perspective of the students' experiences, hopes and feelings put into tension their espoused beliefs about already offering their students high quality opportunities to learn. They began to unpack the challenges of indeed offering opportunities that would foster meaningful learning for all students. To this effect, they were challenged to change their mind frames (Hattie 2012) about their roles as teachers as well as senior leaders.

> Children's comments hurt me. We interviewed two kids who were very clear, very mature to look at things. They said they would like to learn deeper things; that is what they said. Not just dates but to know why something happened, for what purpose, what were the consequences for our present. On the other hand they told us that a good student is one who learns fast, who is connected to the teacher dictating content and able to keep up with the teacher's pace. A good student is the one who can stay at the teachers' pace and a bad student is the one who cannot reach the teacher's speed (Principal, S5, Metropolitan Region, Visit 2).

Working together over the course of 6 months, trust was developed as participants visited each other's schools. Trust enabled deeper reflections with peers that led to self-questioning and the development of new understandings:

> It is very hard for me to get rid of the idea that I must give grades; it is hard for me to give students high grades. I ask students to analyze, to recognize, interpret, to read graphs and I have many failing grades. It is hard for me to give over a 5 [on a grade scale from 1 to 7]. But I want to tell you that I am fighting with myself (Teacher, S4 Metropolitan Region, Visit 3).

> We need students to have an active role in their learning, to strengthen their self-regulation, self-monitoring, and to learn to give peer feedback. The challenge is co-responsibility, to strengthen their autonomy, give responsibility in their learning and to develop higher order abilities (Teacher, S2, Valparaíso Region, Visit 3).

12.7 Participants' Evaluation of SILP

During the final workshop participants were asked to, anonymously, evaluate the usefulness of this professional development program along seven dimensions addressed by the various SLIP components: (a) feedback as a shared topic of inquiry, (b) the review of literature, (c) school visits, (d) data collection tools, (e) theory-engaged evaluation, (f) facilitation provided by the university team, and (g) feedback from peers. Using a scale from 0 (not at all useful) to 10 (extremely useful), the 27 participants rated each dimension and later provided a rationale for the rating assigned. Additionally, they were asked to offer suggestions to improve SILP.

A quantitative analysis shows that, on average, ratings ranged from 10 to 8, thus all dimensions were rated as useful. As shown in quote excerpts indicative of responses provided by a range of participants, the dimension peer feedback (mean rating of 10) was valued because it prompted them to: "*reflect and foster autonomy in making decisions on how to best improve our students' learning*".

Among other dimensions averaging a rating of 9, the chosen topic of feedback was considered highly useful to engage participants in peer review. In the words of one participant, the literature reviewed dimension, "*allowed us to develop a deeper understanding, debunking some myths we had about feedback and learn new categories and ways to provide it. The time to work together on this was insufficient.*"

School visits were also highlighted by all of the participants as a highly useful dimension. This, an infrequent practice among senior school leaders or teachers in Chile, impacted on their beliefs and actions:

> I lived an experience I had not previously had, it opens your mind, the experiences, the strategies that you can use and changes the focus to what really matters which is children, their learning, and how they learn.

Only the dimension relating to the tools for data collection averaged a rating slightly below 8. As shown in the next excerpt, this rating is associated with the differential value assigned to different data collection tools: *"The interviews had a big impact* *in our school as they evidenced our strengths and weaknesses. Observing a class for just 20 minutes is not representative to evidence how much feedback is provided".*

12.8 Discussion and Conclusions

Based on what we learned as we were implementing the different components, and adapting to local conditions, the SILP model evolved from its origin, RiPR (Godfrey and Brown 2019). As we analyzed participants' emergent needs and explicit requests we came to understand that in addition to a network level in which three schools from the same territory collaborate in the peer review process, new learning settings needed to be generated at the school level. With the support from the university team, post visit in-school workshops allowed participants to reflect on how to transfer the learning achieved at the network level into their instructional leadership practices. We found that network level conversations made visible institutional practices, whereas in-school workshops made visible personal theories of actions. Figure 12.2 summarizes outcomes that we have associated with the various learning settings in this professional development model.

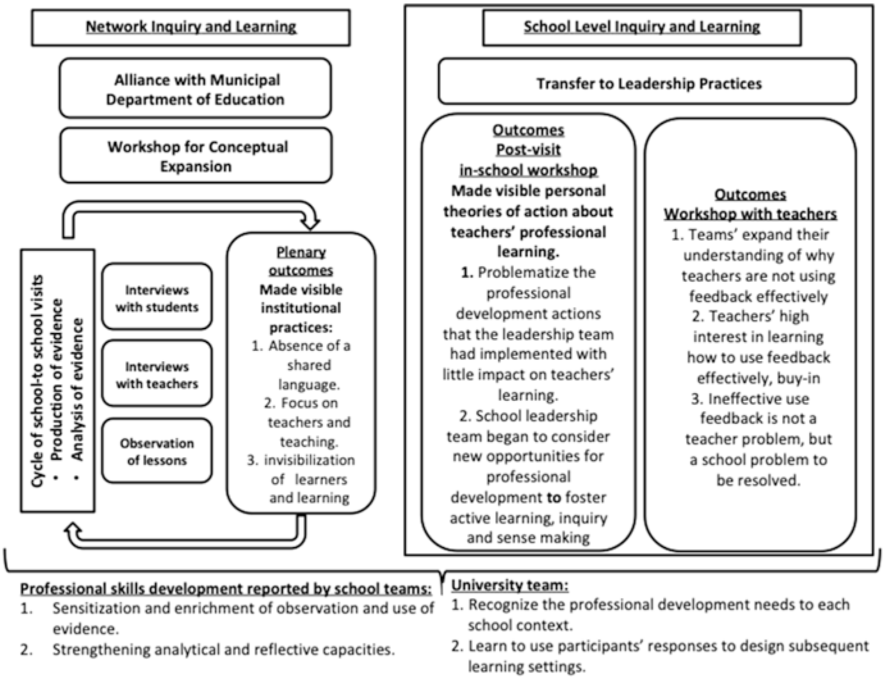

Network Inquiry and Learning
Alliance with Municipal Department of Education
Workshop for Conceptual Expansion
Cycle of school-to school visits
• Production of evidence
• Analysis of evidence
Interviews with students
Interviews with teachers
Observation of lessons
Plenary outcomes
Made visible institutional practices:
1. Absence of a shared language.
2. Focus on teachers and teaching.
3. invisibilization of learners and learning
School Level Inquiry and Learning
Transfer to Leadership Practices
Outcomes Post-visit in-school workshop
Made visible personal theories of action about teachers' professional learning.
1. Problematize the professional development actions that the leadership team had implemented with little impact on teachers' learning.
2. School leadership team began to consider new opportunities for professional development to foster active learning, inquiry and sense making
Outcomes Workshop with teachers
1. Teams' expand their understanding of why teachers are not using feedback effectively
2. Teachers' high interest in learning how to use feedback effectively, buy-in
3. Ineffective use feedback is not a teacher problem, but a school problem to be resolved.
Professional skills development reported by school teams:
1. Sensitization and enrichment of observation and use of evidence.
2. Strengthening analytical and reflective capacities.
University team:
1. Recognize the professional development needs to each school context.
2. Learn to use participants' responses to design subsequent learning settings.

Fig. 12.2 SILP learning settings and associated outcomes

12.9 Network Level Process and Outcomes

This level involves school teams working with other teams in university-based workshops and in school-to-school visits. Although conceptual expansion and refinement processes were present in all of the learning settings, the design identifies distinct implementation milestones. The first pre-school visits workshop conducted at the university in each region provided a general framework to enrich the theoretical understanding of feedback and what needed to be mobilized (theories of action) to change practices. The second pre-school visits workshop sought to refine observation skills and understanding of the tools used to produce evidence during the school visits. The final workshop, taking place after all school visits had been completed, involved a metacognitive exercise to consolidate learning. For example, participants made explicit changes in their notions of learning. Additionally, questions of equity in the opportunities to learn afforded to different students were now made visible when conducting classroom observations.

School-to-school visits enabled the use of evidence from different sources to construct an overview of the school regarding a specific pedagogical practice - in this case feedback in the classroom – and institutional practices that needed changes. For example, they identified a need to change faculty meetings in which conversations were mostly about controlling students' misbehaviors and hardly about enhancing opportunities to learn. Building a new vision together with the other schools in the network enabled a broader and deeper perspective on how teaching and learning were enacted at the school. Key leverage points were identified as it became clear that student's own expectations for deep learning were not being met through the lessons teachers were delivering.

Across all nine schools, the absence of a shared pedagogical language and a vision of learning that has students as protagonists emerged as a challenge to be worked on. Participants became aware of how feedback could promote the learning of all students, moving forward their spouse moral imperative to advance social justice by understanding that powerful learning was a tool for social transformation in the lives of their students. Participants also became aware that the instructional core had become heavily focused on meeting external accountability demands, leading to the dearth of conversations about teaching and learning at the core of their leadership practices.

12.10 School Level Process and Outcomes

The network level learning settings described so far for SILP addressed the "what" of the change process, but not "how" to transfer this learning into these participants' instructional leadership practices. Developing additional learning settings to support transfer was key to the overall purpose of this PD program. When schools were invited, an explicit expectation was that peer review findings would be used to enhance effective feedback to support students' learning.

During the post-visit in school workshops school leaders' theories of action about teachers' professional learning were the object of critical reflection. When we met with the host school team, we questioned their impetus to immediately initiate actions to address the problems identified in teaching and learning. The teams initially blamed teachers for the instructional practices present in their schools, claimed they were resistant to change, and resorted to implementing new teacher supervision measures, bringing in external experts to lead teacher training, and placing additional requirements on teachers. The actions they were proposing were the same they had already been implementing with little impact in terms of helping teachers develop an enriched pedagogical repertoire. Once they became aware that these solutions had not been useful in the past, that their own learning through SILP had followed a different set of assumptions about professional learning, they began to think of new possible forms of professional development opportunities they could afford to the teaching staff.

These leadership teams had major concerns about how to best share the peer review findings with teachers. Moreover, they explicitly requested the assistance of the university teams on this matter. As we planned with them the workshop to share the peer review findings we were explicit about how active learning, inquiry and sense making for the evidence produced could be fostered or hindered. The participation of teachers in these conversations was key to breakdown stereotypical perspectives that among senior leaders that tended to position teachers as the problem versus problematizing the fit between leadership practices and teacher's identities, needs and expectations. These new insights can become highly relevant to meet demands from the National System for Teachers' Professional Development that require principals to develop local professional development plans that foster professional learning communities within their schools.

In conclusion, this PD model allowed for the construction of a "third space" (Zeichner 2010), where academic knowledge and practical knowledge came together to problematize theories of actions that could lead to new practices for all involved. After the completion of the cycle of school-to-school visits we realized that it is not enough for senior leaders to develop new conceptual tools and data production tools to make learning visible at their schools. They also needed to learn how to deploy these new tools through their instructional leadership practices in order to enable changes in how teachers provided feedback to students. Therefore, a professional development model for these Chilean school leader needed to attend more explicitly to this aspect, something not as prominent in the RiPP model developed for the UK context. As researchers and professional developers worldwide look for innovative ideas to support school improvement, careful attention and responsiveness to the learning needs of participants is key to adapt rather than adopt what has shown to work elsewhere.

References

Ahumada-Figueroa, L., González, A., & Pino-Yancovic, M. (2016). *Redes de Mejoramiento Escolar: ¿Por qué son importantes y cómo las apoyamos?* Informe técnico n°1. Valparaíso, Chile: LIDERES EDUCATIVOS, Centro de Liderazgo para la Mejora Escolar.

Bolivar, A. (2015). Un liderazgo pedagógico en una comunidad que aprende. *Padres y maestros, 361*, 23–27. https://doi.org/10.14422/pym.i361.y2015.004.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3–15.

Carrasco, A., Gutiérrez, G., & Flores, C. (2017). Failed regulations and school composition: Selective admission practices in Chilean primary schools. *Journal of Education Policy, 32*(5), 642–672. https://doi.org/10.1080/02680939.2017.1312549.

Chapman, C., & Muijs, D. (2014). Does school-to-school collaboration promote school improvement? A study of the impact of school federations on student outcomes. *School Effectiveness and School Improvement, 25*(3), 351–393.

Chapman, C., Ainscow, M., Bragg, J., Gunter, H., Hull, J., Mongon, D., Muijs, D., & West, M. (2008). *Emerging patterns of school leadership: Current trends and future directions*. Nottingham: NCSL.

Cortez, M., Campos, F., Guzmán, J., Rojas, J, Ulloa, J., Gajardo, J., Peña, M., Albornoz, C., Montecinos C., Godfrey, D. (2019). Indagar colaborativamente sobrer etroalimentación efectiva en aula para potenciar liderazgo pedagógico de equipos directivos. Paper presented at the *III Congreso Internacional Sobre Liderazgo y Mejora Escolar, CILME 2019.* Santiago.

Godfrey, D., & Brown, C. (2019). Innovative models that bridge the research–practice divide: Research learning communities and research-informed peer review. In An ecosystem for research-engaged schools (pp. 91-107). London: Routledge.

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Abingdon: Routledge.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. London: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.

Kalinowski, E., Gronostaj, A., & Vock, M. (January-March 2019). Effective professional development for teachers to foster students' academic language proficiency across the curriculum: A systematic review. *AERA Open, 5*(1), 1–23. https://doi.org/10.1177/2332858419828691.

Lave, J. (1988). *Cognition in practice*. Cambridge: Cambridge University Press.

Montecinos, C., & Cortez, M. (2015). Experiencias de desarrollo y aprendizaje profesional entre pares en Chile: implicaciones para el diseño de una política de desarrollo docente. *Docencia, 55*, 52–61.

Montecinos, C., Ahumada, L., Galdames, S., Campos, F., & Leiva, M. V. (2015). Targets, threats and (dis)trust: The managerial troika for public school principals in Chile. *Education Policy Analysis Archives, 23*(87). https://doi.org/10.14507/epaa.v23.2083.

OCDE. (2017). *Evaluación de Políticas Nacionales de Educación. Educación en Chile.* Organización para la Cooperación y el Desarrollo Económicos (OCDE).

Robinson, V., Hohepa, M., & Lloyd, C. (2009). *School leadership and student outcomes: Identifying what works and why: Best evidence synthesis iteration*. Wellington: New Zealand Ministry of Education.

Säljö, R. (2009). Learning, theories of learning, and units of analysis in research. *Educational Psychologist, 44*(3), 202–208.

Stake, R. (2010). *Qualitative research: Studying how things work*. New York: Guilford Press.

Valenzuela, J. P., & Montecinos, C. (2017). Structural reform and equity in Chilean schools. *Oxford Research Encyclopedia of Education*. https://doi.org/10.1093/acrefore/9780190264093.013.108.

Valenzuela, J. P., Bellei, C., & De los Rios, D. (2014). Socioeconomic school segregation in a market-oriented educational system. The case of Chile. *Journal of Education Policy, 29*(2), 217-241. https://doi.org/10.1080/02680939.2013.806995.

Zeichner, K. (2010). Rethinking the connections between campus courses and field experiences in college and university- based teacher education. *Educação Santa Maria, 35*(3), 479–501.

Mónica Cortez is coordinator of professional development at the Centro Líderes Educativos PUCV at the Pontificia Universidad Católica de Valparaíso. Her most recent co-authored publications, related to the subject of this book are: Montecinos, C., Cortez, M. and Walker, H. (2015) "School administrators' understandings and management of barriers for the school's involvement in the practicum component of initial teacher education in Chile", in*International Journal of Educational Development*; Cortez, M. and Montecinos, C. (2016). Transitando desde la observación a la acción pedagógica en la práctica inicial: Aprender a enseñar con foco en el aprendizaje del alumnado [Moving from observation to pedagogical actions in initial field experiences: learning to teach with a focus on students' learning], in *Estudios Pedagógicos*.

Fabián Campos is executive secretary at the Centro Líderes Educativos PUCV at the Pontificia Universidad Católica de Valparaíso. His recent co-authored publications include: Learning to Lead the Change in School Organizations in Chile (2019), *Magis. Revista Internacional de Investigación en Educación*; Novice principals in Chile mobilising change for the first time: challenges and opportunities associated with a school's readiness for change (2018), *Educational Management Administration & Leadership*; An analysis of the use of relational expertise, relational agency and common knowledge among newly appointed principals in Chile's public schools (2017), in Edwards, A., (Ed.), *Working Relationally In and Across Practices*.

Carmen Montecinos is a professor in the School of Psychology and director of the Centro Líderes Educativos PUCV at the Pontificia Universidad Católica de Valparaíso, and a senior researcher at the Center for Advanced Research in Education (CIAE-Universidad de Chile). She has published widely in the field of educational leadership and among her most recent co-authored publications are: Bush, T. and Montecinos, C. (2019). Leadership preparation and development, in Bush, T., Bell, L. and Middlewood, D. (Eds.), *The Principles of Educational Leadership and Management* and Moving the school forward: problems reported by novice and experienced principals during a succession process in Chile (2018), *International Journal of Educational Development.*

Jorge Rojas is an assistant professor in the College of Education at the Universidad de Concepción, and working in the Centro de Liderazgo para la Mejora Escolar in the Bio Bio Region of Chile. He has worked on a project titled School leadership and accountability in Chile and his most recent publications include: Gwynne, K. Rojas, J., Hines, M., Bulkeley, K., Irving, M., McCowen, D., and Lincoln, M., (in press). Customised approaches to vocational education can dramatically improve completion rates of Australian Aboriginal students.*Australian Health Review.*

Marcela Peña is coordinator of the area of Innovation and Development of the Centro de Estudios y Desarrollo de Educación Continua para el Magisterio, Saberes Docentes, of the Universidad de Chile, and professor at the Universidad Academia de Humanismo Cristiano. Her publications include the collective work developed by the Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas (CPEIP), in conjunction with Harvard University through The Harvard-Chile Innovation Initiative and David Rockefeller Center for Latin American Studies, *Facing Challenges: Case studies for school directors* (2015).

Jorge Gajardo is an assistant professor at the College of Education at the Universidad de Concepción and Program Coordinator in the area of capacity building for school leadership of the Centro de Liderazgo para la Mejora Escolar. He has collaborated as author and co-researcher in publications in the area of school leadership and school climate. His most recent co-authored publications include: Nail, O., Gajardo, J., Valdivia, J., & Norambuena, D. (2019) Experience in training teachers in initial training for improving school climate and preventing school violence, in Torres, B (Eds.) *Education, School and Teachers*.

Jorge Ulloa is an Associate Professor of the Department of Educational Sciences and Director of the Master in Educational Management and Leadership of the College of Education of the University of Concepción. He is also Head of the School Leadership Development Line of the Centro de Liderazgo, Lideres Educativos. His recent publications include Ulloa, J. (2013) Strengthening School Leadership: A Contribution from the Critical Incident Analysis Technique, in O. Nail (2013), *Analysis of Critical Classroom Incidents: A tool for improving coexistence*, Santiago de Chile: RIL Editores and Ulloa, J. and Rodríguez, S. (Eds). (2014) *School leadership and professional development of teachers. Contributions to school improvement*. Santiago de Chile: RIL Editores.

Charles Albornoz is a teacher, Postgraduate in Curriculum and Assessment, Master in Management and Educational Leadership and Doctor in Education. He has served as school principal and as an advisor to the Ministry of Education in the design and implementation of training programs for pedagogical technical advisors and member of the Curriculum and Evaluation Unit of the aforementioned ministry. He is currently a professional in Saberes Docentes - Center of Studies and Development of Continuing Education for Teaching in Universidad de Chile and Territorial Coordinator in the Leadership Development Area within Primary Schools and High Schools at the Centro de Liderazgo para la Mejora Escolar.

Part VII
Synthesis and Discussion

Chapter 13
Evaluation Theory and Peer Review. Practice, Policy and Research Implications

David Godfrey

Abstract This chapter compares and contrasts ten models of peer review and six countries described in the earlier chapters of this book, using a conceptual framework that looks at use, methodology and values underlying their evaluation theories. Following this analysis, the chapter outlines how peer review plays a particularly strong role in the leadership development of those who participate in them. The issue of trust as a mediating factor for peer review participation and success is explored. The ability of peer review cooperation to build collective efficacy and therefore increase student attainment is also briefly outlined. Finally, different network level configurations that surround peer reviews, and the effects of these variations, are explored. The findings have implications for policy makers; school leaders and researchers working in this field.

13.1 Introduction

This chapter brings out some lessons for theory, practice and policy from the peer review programmes mentioned in the preceding chapters. Ten distinct models have been described in this volume, in the contexts of 6 countries, namely: England, Wales, Australia, Bulgaria, The Czech Republic and Chile. The place of peer review within these school systems is described throughout, along with the specific aims of each programme, evidence for their benefits and the challenges of building and sustaining peer review networks. In doing so, this book adds considerably to the limited evidence base that exists for school based peer review worldwide. Further analysis follows in this chapter, synthesizing some of the key learning from this illuminating collection of case studies.

The chapter starts with an analysis of the evaluation principles underlying each approach, as explicitly stated by the authors or creators of the programmes or by an implicit analysis of the methods and values described within each one. The

---

D. Godfrey (✉)
UCL Institute of Education, London, UK
e-mail: david.godfrey@ucl.ac.uk

© Springer Nature Switzerland AG 2020

D. Godfrey (ed.), *School Peer Review for Educational Improvement and Accountability*, Accountability and Educational Improvement,
https://doi.org/10.1007/978-3-030-48130-8_13

conceptual framework of Christie and Alkin (2012) is applied to this task. This framework allows for a comparative analysis on the three 'branches' of evaluation theory: use, methodology and values.

Following this analysis, the chapter gives a wide-ranging discussion of the benefits, impact, mediating and contextual variables involved in the implementation of peer review. These have implications for policy makers wishing to bring in peer review as system strategy; school leaders and other practitioners wishing to involve themselves in a peer review programme – or set up their own; and to researchers working in this field.

13.2 A Conceptual Framework for Peer Review

Christie and Alkin (2012) have used the metaphor of a tree to model the origins of various strands of evaluation activity. At the roots of the tree, are social *accountability, social inquiry* and *epistemology*. The first of these is described as a consistent theme in the motivation for conducting evaluation of programmes. As described in Chap. 1, and indeed throughout this book, peer evaluation is seen to fill a needed gap in professional, moral and lateral forms of accountability, in contrast to the market and contractual forms of accountability that dominate in external evaluation. Social inquiry is the research origin of evaluation, in that it involves a systematic process of discovery and learning and sharing this with others. The epistemological level regards the way in which knowledge about the phenomena is to be understood and the relationship of this knowledge to the evaluator. Above these roots are the corresponding branches of how the evaluation theory relates *to use (and users*), the *methodologies* employed during the evaluation (research instruments and processes) and the ways in which judgements are made and how the evaluand is *valued*. These are further elaborated below in order: use, values then methods for a more logical flow in the analysis of each peer review model.

13.2.1 Use and Users

Theorists concerned with *evaluation use* emphasise how evaluation can and should inform decision-making. This particularly relates to how evaluations lead to organisational learning and to increased agency of stakeholders in the process of change. Theorists such as Fetterman (1994) and Cousins and Earl (1992) fall within this branch. The former describes his approach as empowerment evaluation and is described in Chap. 10. In this model of evaluation, the primary goal is for the participants to gain self-determination through the process of evaluation. Therefore, stakeholders design their own performance indicators, evaluation criteria, focus, evaluation instruments and methods. Coaches or facilitators do not run the evaluation but guide stakeholders in the process of constructing their own. Evaluation use is designed to emancipate stakeholders by illuminating and demystifying their practices. The process is designed to demystify the evaluation process, and emancipate stakeholders through a process of illumination about their practices.

In participatory evaluation (Cousins and Earl 1992) collaboration occurs between evaluators (or researchers) and practitioners; it is concerned with the inclusion of a wide range of stakeholders in consultation with the evaluation process. Important to this type of evaluation is to consider who is involved, particularly as 'primary stakeholders'. In the analysis of peer review programmes, this concerns which personnel (within and outside of schools) are involved in the programme. It is also relevant to understand the organisational capacity of the participating schools, as this would affect utilization of the evaluation. The inspection 'categories' of schools can be taken as proxy indicators of school capacity in this regard. The ways in which leadership of evaluation and evaluation use is encouraged and the locus of this leadership is a further area of analysis.

13.2.2 Values

This branch of evaluation is concerned with the distinction between facts and values, and evaluation as an act of judging the value of something. This branch is strongly influenced by epistemology; i.e. the extent to which judgements are trying to achieve objectivity versus those that are subjective and more firmly rooted in the 'social constructivist' research tradition. The latter sees reality as rooted in the perspectives of various actors and is more clearly linked to the particular context and circumstances. Evaluation theorists such as Eisner (1985) reject objectivist and quantitative approaches to evaluations, focusing more on the role of the connoisseur. A connoisseur, by virtue of training, study and experience, is better able to make subtle distinctions and thus judgements about the quality of what they observe. In peer reviews, reviewers can act as arbiters of school quality, however the degree to which judgements are negotiated with the host school in a review varies between one peer review programme and another. An objectivist stance on valuing would lead to a higher degree of consistency between various evaluators. The principal trade off is between the extent to which evaluation findings can be compared and generalized across settings versus the extent to which evaluations have greater 'authenticity' and detail in relation to a particular programme. In the case of our peer reviews, the criteria used in framing the reviews and the types of categories or scales used (or not used) are relevant. Some peer reviews for instance, chose a scale that reflects the external accountability framework while others consciously do not. Those that use inspection criteria may have greater alignment and build more on past or future external evaluations while those that are different (or tailored to each review) may have more of a formative intention and/or involve aspects not covered by inspections.

13.2.3 Methodology

The evolution of programme evaluation theories mirrors that in the social sciences overall. As with the above 'values' branch, this is essentially a contrast between positivist and post-positivist vs. interpretivist perspectives. The former emphasise rigour, collecting data from highly controlled circumstances, determining impact in relation to control or comparison groups. Proponents such as Campbell and Stanley (2015) have proposed experimental methods, for example randomized controlled trials in this pursuit of generalizable results and comparability. Others, like Cronbach and Shapiro (1982), have suggested the use of statistical tools to find causative relations between events and to establish construct validity. By contrast, other evaluation scientists, such as Ralph Tyler (e.g. Tyler 1942), reject the use of quantitative data; proposing an objective-oriented approach. This approach involves categorising the objectives and collecting behavioural evidence of their occurrence using appropriate, trialed instruments for measurement. This kind of thinking has dominated the educational field and is particularly notable in the methods used in school inspections. Moving further along this continuum are approaches to enquiry that are akin to collaborative action research (Carr and Kemmis 2003), involving cyclical iterations of learning or those requiring rich detail and contextual analysis, as in case study research (Stake 1995).

Analysis on this branch looks at evaluation practices in peer reviews; particularly how the focus for reviews is arrived at and then how the data in peer reviews are collected, analysed and reported.

13.3 A Comparative Analysis of Peer Review Programmes

Below is the application of the above categories to a comparative analysis of the 10 models of peer review described in the preceding chapters of this book. A summary of this analysis is shown below in Table 13.1.

13.3.1 Use/Users

All of these peer review models have a strong emphasis on use and therefore draw on the work of theorists who are concerned with, "direct program site use (in action or understanding) that results from a particular evaluation study" (Christie and Alkin 2012, p. 1). The strong formative emphasis of peer review means that both visiting reviewers and host schools seek to gain further understanding of the issues (however defined) that they focus on and to use these to propose improvements.

Each school is able to select its own area of focus albeit these are constrained in a number of differing ways. For instance, the participants in Research-informed

Table 13.1 summary of evaluation characteristics of each peer review programme

Peer review programme and its aims	Use/Users (Social accountability)	Valuing (Epistemology)	Methodology (Social enquiry)
Chapter 2: Queensland: Aim: To complement system reviews, feedback on improvement trajectory	Schools: Earned autonomy for schools with good improvement trajectory.	Objectivist: Nine domains of the *National School Improvement Tool* (NSIT) used. School improvement unit (SIU) trained reviewers.	Reviewers use system data (such as standardised test results, report card results, school community satisfaction, attendance and behaviour data) and undertake extensive fieldwork at the school, including interviews with school leaders, staff, students, parents and community members. School self-evaluation/reflection also used.
	Review teams: SIU reviewer, a contracted reviewer (selected from a pre-qualified panel) and a principal peer reviewer. Deputy principals and heads of department, and other school staff can be trained in SI.	The review team co-creates the review report and presents findings and improvement strategies, organised against the nine domains of the NSIT	
Chapter 3: Peer enquiry, central South Wales	Schools: Initially those categorized as ‘green’. Later yellow and amber.	Subjectivist: Schools work in joint practice development model – Mutual learning.	Collaborative enquiry: Review team follow agreed ‘lines of enquiry’. These determine the data collection.
Aim: To replace previous local SI infrastructure	External ‘challenge advisers’ (CAs) invited to feedback and to help broker support. CA has more role in focus of review for Amber schools.	However, focus for amber schools more on meeting specific needs of groups of children or following an external inspection.	
	Lead peer enquirer (LPE), a Headteacher with proven leadership expertise, a second serving Headteacher and an associate peer enquirer - an individual seen be aspiring to headship and affiliated to the LPE or second serving Headteacher.		

(continued)

Table 13.1 (continued)

Peer review programme and its aims	Use/Users (Social accountability)	Valuing (Epistemology)	Methodology (Social enquiry)
Chapter 5: England. Instead reviews (NAHT)	Schools categorized as ‘good’ or ‘outstanding’.	‘Soft’ objectivist: Framework developed by NAHT in consultation with members and other experts. Not graded, rather what should be prioritised, developed, maximised or sustained. Wider focus than external framework, including ‘school vision’ for instance.	Two-day school visit by 2 principals and the Lead reviewer. Reviewers collect data according to focus of the review. Report is written by Lead reviewer and findings/ conclusions negotiated with school and other reviewers.
Aim: To assist Headteachers in their school self-evaluations. To share expertise between current primary school head teachers.	Head teachers work in self-formed clusters. One Lead reviewer, usually experienced in inspection methodology from outside the cluster.		
Chapter 6: Peer review in Bulgaria in polycentric model	Principals in a regional school network. Later inspectors validate/ challenge network according to aims of the reviews. Strongly user oriented at the network level.	Constructivist: Thematic framework (parental involvement) developed by schools, regional inspectors and researchers. Four standards for quality parental involvement in school.	Four data collecting instruments for gathering self-evaluation data, from parents, teachers, school leaders, students.
Aim: To build a polycentric model of evaluation of a network of schools.			Co-construction of data collection tools, identification of evidence, and the report format
Chapter 11: RiPR	Head teacher, senior leader, middle leader and teacher. Clusters self-formed. Any category could take part but so far none in ‘requiring-improvement’ category has participated. University facilitation. Schools given input in workshops on implementation based on research	Constructivist: Theory-engaged evaluation model – Emphasises need to recognize values of teachers. Not based on external inspection criteria.	Enquiry based methodology introduced by university facilitation guides co-construction of data collection tools. Evidence-based classroom observation categories. Policy documents analysed for espoused theory for action. University facilitator guides learning discussion
Aims: To work collaboratively to improve feedback and assessment policies in a cluster of schools			

(continued)

Table 13.1 (continued)

Peer review programme and its aims	Use/Users (Social accountability)	Valuing (Epistemology)	Methodology (Social enquiry)
Chapter 12: SILP (Chile)	School principal, plus SLT, middle leaders, teachers school clusters identified by a leadership Centre as good 'early adopters'. Middle/low middle and insufficient categories. University facilitation. Schools given input in workshops on implementation based on research.	Constructivist: Theory-engaged evaluation model – Emphasises need to recognize values of teachers. Not based on external inspection criteria.	Enquiry based methodology introduced by university facilitation policy documents analysed for espoused theory for action. More prescription from facilitators about process but some latitude on interview questions to review team. Follow up visits to school by university team to encourage leadership to act on findings and provide support.
Aims: As above			
Chapter 7: Czech peer review "road to quality' project.	Schools: Not dependent on inspection grade.	Objectivist: Initial SSE based on six areas of quality as set out in Czech legislation. Common European framework for peer review with 14 areas of quality, adapted to Czech setting, based on vocational model (Gutknecht-Gmeiner 2007). Visits were mostly reciprocal.	Visits to classrooms, interviews with teachers and students and SSE and other data examined. Report sent by reviewing team and school could suggest amendments. No set use of methods or tools. Process depended on competencies and experiences of reviewers.
Aim: To build a system of support to increase the capacity to conduct self-review	Review team: School principals, students, institutions, administration bodies, partners in vocational education and training, parents and prospective employers. Schools encouraged to implement recommendations. External facilitator advises school and reviewing team.	Descriptive feedback rather than judgements.	

(continued)

Table 13.1 (continued)

Peer review programme and its aims	Use/Users (Social accountability)	Valuing (Epistemology)	Methodology (Social enquiry)
Chapter 10: NSW QA process.	Schools: Not dependent on inspection grade.	Constructivist (group 1): Development of own exemplary practice statements by (group 1) principals.	Group 1 principals were guided to create their own evaluation tools by critical friend. Group 2 used tools in as part of reviewing team. Guided in following an empowerment evaluation approach.
Aim: Robust, useful, and influential evaluation of whole-school performance and governance for the school being reviewed. Building evaluation capacity and trust	Principals decided on composition of the team according to school focus. Aided by trained and qualified critical friend (Region's senior school accountability and improvement officer).	'Soft objectivist' (group 2) participatory and collaborative evaluation but not involved in designing criteria.	
Chapter 8: Challenge partners (CP) quality assurance	Schools: All grades participate; organized into hubs around teaching schools alliances.	Objectivist: Uses existing Ofsted inspection framework as the basis of review and. School areas graded as outstanding, good, requires improvement, insufficient.	CP work with hub schools to coordinate review visits and training. Pre-review conducted by reviewing team based on public data. School provides improvement plan. One and a half day review visit. Joint lesson observations and meetings with pupils and key staff. Report produced collaboratively and shared internally and externally. Expectation that findings incorporated into school development plans.
Aim: Validates the school's improvement action plan.	Senior leadership of host school receive review visit conducted by team of reviewers outside of immediate vicinity. Review team consists of 2–4 senior leaders and led by experienced Ofsted trained reviewer. Includes validating areas of excellence and accrediting school to lead workshops/training in this area, recorded in subject directory.		
Accredits areas of strength that can be shared within the hubs and across the partnership.			
Identifies areas for development			

(continued)

Table 13.1 (continued)

Peer review programme and its aims	Use/Users (Social accountability)	Valuing (Epistemology)	Methodology (Social enquiry)
Chapter 9: Schools partnership programme	Schools: All Ofsted grades can participate	Soft objectivist: Methods used were co-constructed with school leaders over time. Self-evaluation framework with enquiry questions that focus on development. Not based on external inspection framework. Learning assumed to occur through reviewing as well as by being reviewed. Non-judgemental approach to reviewing is emphasised ('open learning conversations'). Credibility of reviewers comes from skill in reviewing not performance of their schools.	Lead reviewer has pre-review meeting with host principal to agree focus for review. Reviewers encourage to use range of data, including 'small data' (teacher's reflections) as well as standardised public data. Review team feeds back to host school senior leaders at end of day and brief report is sent to school. Later, ICs facilitate improvement workshop according to the recommendations of the review. Workshop may draw on published academic research.
Aims: To provide validation and challenge for school's self-review and to identify and encourage individual school improvement and school to school support. Aims to develop culture of shared responsibility and trust (openness).	Strong developmental orientation based on appreciative enquiry approach.		
	Schools work in clusters (self-formed and/or existing) for reciprocal reviews. One head teacher coordinates partnership. Head teachers lead reviews, senior leaders also take part in reviews and middle leaders trained as improvement champions (ICs).		

Peer Revew (RiPR) and Schools Inquiring and Learning with Peers (SILP) models agree to work on a particular theme but within that they can choose their narrower aims for the evaluation. In other models, such as the Schools Partnership Programme (SPP), there is a framework for self-evaluation but schools then derive their own specific enquiry questions for the team to review. In the Queensland review model the external inspection framework is used, and schools can either ask for a whole school 'scoping' review or to focus on particular areas. These reviews serve as 'interim' evaluations when schools are a few years away from a cycle of external evaluation. Peer evaluations often seek to build on the successes of schools too, adopting an appreciative inquiry approach (Cooperrider and Srivastva 2017). For instance, the Instead model allows for 'celebration' of good practice in the report. This contrasts with how external inspections are generally conducted, where a deficit model is used in which evaluators' recommendations focus on improvements to the weakest areas.

Aspects of empowerment evaluation approaches are evident throughout, most explicitly in the New South Wales quality assurance process. Ikin's description (Chap. 9) of principals in group 1 particularly illustrates this. Evaluators are the school principals themselves who decide the framework, the criteria to use and the evaluation tools. This embodies Fetterman's evaluative approach:

> Empowerment evaluation is explicitly designed to serve a vested interest – program participants. It is designed to help them become self-determined … As they begin to engage in the self-evaluation process, they quickly learn how to be systematically analytical about themselves and their program … They also become accustomed to justifying and documenting the basis for their assessments. (Fetterman 1994, p.10)

Thus the development of evaluation capacity is seen as an end in itself in the NSW peer review programme. In order to support capacity building for evaluation, many of the programmes mentioned above employ someone from outside of the immediate stakeholder group to assist in the design and implementation of the peer review. These outsiders are sometimes trained inspectors, such as in the Challenge Partners (CP), Instead and Queensland models or specialists in evaluation (RiPR/SILP). The Czech model is unique in using stakeholders outside of the school sector to assist in evaluations.

A strong element seen across the above peer review models is the ways in which they foster leadership. In the participatory model of evaluation, this is about encouraging the development of leaders who can help support the evaluation process and then drive through the implementation of the recommendations. In the case of the RiPR and SILP models, these models consciously involve aspiring leaders, or informal leaders with high practice based social capital. These are often credible teachers known to be excellent practitioners who others turn to for advice, sometimes known for being early and enthusiastic adopters of innovations or frequent users of research. The SPP model employs middle leaders in facilitating change in one of the partner schools in their cluster. The coaching model that these improvement champions employ is seen as a key ingredient of its success in developing system leaderships and building agency for change. The leadership development potential of peer reviews is discussed further below.

A final dimension to consider is the type of schools involved in these peer reviews as determined by their current improvement trajectory. If categorization by an external inspectorate is taken as a reasonable proxy, then we can see a number of programmes where schools tend to be comprised of mostly ‘good or excellent’ (or the national equivalent) in their category. The implication may be that lower category schools may have less capacity to take part in peer review as they are focusing on specific improvement targets recommended in an inspection report or that they may feel they have less to offer in terms of good practice to higher performing partners. In the Instead programme lower performing schools did not take part and the RiPR peer reviews, it has been the case that lonely good and outstanding category schools have taken part (so far, through self-selection). In the case of Queensland, schools on a good improvement trajectory were able to voluntarily opt in to the peer review cycle. This is a good example of earned autonomy, where schools deemed to have shown the ability to improve are given more latitude to rely on their own self-evaluations to continue to improve (Suggett 2015). However, we do see examples of peer review programmes where schools on lower categories work alongside those on higher ones, such as in the SILP example in Chile. Similarly, the CP programme opts for the logic of the London Challenge for schools with higher capacity to help lower performing schools in the same network. In the case of Wales, the peer review programme started with schools in the higher green and then amber categories, but once established then brought in more ‘needy’ ‘red’ schools. For the latter schools, the reviews became more targeted, aligned to immediate priorities as stated in recent inspection reports. Where schools are clustered, it would seem logical that if all the schools are lower performing, then a peer review process, however well facilitated, would struggle to work. Such schools may need further training and external expertise for leadership and teachers in order to make the strides needed. The ways in which all these different dynamics play out remains an interesting source of future analysis.

13.3.2 Values

One of they key aspects to analyse in peer review is the type of evaluation framework used by reviewing teams. The programmes mentioned here show a number of variations in approach, indicating either an objectivist or a subjectivist/constructivist stance (Fig. 13.1).

For each of the above, there is a continuum rather than a clear-cut or dichotomous relationship. In the first variation, CP uses the Ofsted inspection framework for its reviews, thus adopting a more objectivist approach. With a trained reviewer, this model should give the school a good idea of how they would perform if formally inspected. In contrast, the empowerment evaluation approach used in the New South Wales model is more constructivist, as shown through stakeholders’ construction of exemplary practice statements. While not providing the assurances offered in the CP model, this framework allows for developmental reviews using criteria that reflect the values of stakeholders themselves.

For the second distinction, the Queensland model exemplifies the broad and standardized approach using a framework that covers the nine domains of the National School Improvement Tool. This looks at areas as diverse as: the school’s culture of learning, its community partnerships and pedagogical practices. At the other end of

Variation	Objectivist	Subjectivist/Constructivist
1	Framework used by an external inspection	Constructed by evaluators/stakeholders
2	Broad and standardised across all reviews	Customised for each review focus
3	Summative grading	Formative levels or descriptive
4	Focused on assessing the school	Reciprocity/mutual learning

Fig. 13.1 Variations in valuing approaches in peer review frameworks

the continuum, the thematic model employed in Bulgaria was constructed specifically to look at parental involvement and was customized for this purpose.

Moving to the summative/formative balance, the CP framework employs grading (Ofsted levels), albeit for use formatively, while the Instead model uses self-evaluative statements (prioritized, developed, maximized or sustained). The latter avoids grades, suggesting in place the stage of development for the school on a particular dimension. Further along this continuum, RiPR/SILP models are more descriptive; reviewers collect data based on a research focus, such as teachers' uses of feedback in the classroom.

Finally, the focus on assessing the school is more evident in the CP model. CP uses reviewers who come from outside the immediate local environs of the school and this is designed to allow reviewers to have a greater sense of impartiality when making judgements. By contrast, the cluster approaches of models like Instead and SPP involve reciprocal review visits, each school taking turns to review and be reviewed. Mutual learning is most evident in the SILP/RiPR models due to the emphasis on a joint practice development model. Here, successive visits are designed to deepen all participants understanding of a discrete area of practice.

13.3.3 Methodology

None of the peer review programmes adopts a strong positivist scientific orientation to gathering data, rather, they take approaches from applied social science. Thus statistical methods of validity are not considered since the uses and values dimensions of evaluation are paramount, especially the former. In other words, the aims of the peer reviews are to inform organisational learning and development rather than to provide statistically sound comparisons between schools or teachers. Nevertheless, the large peer review programmes are now in a position to provide 'meta-reviews'. In the case of the SPP, each partnership conducts a so-called 'review of reviews' to look at shared learning. These inform school-to-school support and strengthen the notion of collective responsibility for bringing about improvements in the cluster. The presence of hundreds of schools in such programmes creates further opportunities for wider comparisons and learning across the system.

The basic practices of peer review: observing classrooms, scrutinizing student work and interviewing students, staff and other stakeholders, are common across all the models. Most peer reviews are likely to make reference to school data at the pupil level for progress and attainment too (although the extent to which national benchmark data available varies from system to system). In this respect peer reviews are very similar to methods used by inspection teams, albeit the stakes are lower, and reports are generated for internal use rather than for public scrutiny. Nevertheless, a degree of professional pride means that the review process can be quite stressful for staff involved, as Chap. 5 shows.

A further distinction between all the peer review programmes, and external evaluations is the power that the host school has to focus the enquiry on aspects of their own choice. Although this is partly shaped by a review framework or theme, the school can steer the direction of the review by setting an enquiry question (as in the SPP), co-constructing data collection instruments (e.g. RiPR) and proposing the timetable for the review and personnel to be interviewed (all models). The agenda for peer review is thus trust based; reviewers take for granted that the host school wants them to see all the evidence needed to conduct an accurate review on the focus its leaders have chosen.

The RiPR and SILP models are the most methodologically aligned as a form of action research. Although documentary analysis forms part of most reviews, RiPR is perhaps the most distinctive in that schools' policy documents are examined using a theory of action analysis of intended aims, actions and consequences. Furthermore, the theory-engagement process of evaluation (Robinson and Timperley 2013) is cyclical, starting with the identification of a problem and working through to an agreed evaluation of the success of an alternative approach. Following Argyris' (1976) double-loop learning concept, the next iteration begins with a re-defined problem to resolve. Some length is dedicated in these models to discussion and collective learning from the data analysis; the written report itself is very brief, and less important than this dialogue. This collective enquiry approach means that the facilitation by university staff is an essential part of the process in these programmes, in contrast with others that tend to use trained inspectors as lead reviewers. Facilitators in RiPR and SILP have additional roles as 'critical friends' – as used in action research approaches - who provide an outsider's perspective, link analysis to published research and challenge insiders' interpretations of evidence.

The above evaluation theory conceptual framework has provided a more thorough comparative analysis than conducted previously in the literature on peer review. The findings here reflect the spectrum of evaluation intention illustrated in Table 13.1, i.e. peer reviews are oriented more towards professional than external accountability, and more growth than control oriented, particularly when compared with external evaluations or inspections. With this orientation, trust becomes a key variable in the effectiveness of peer review (see discussion section) whereas in external evaluations the inspectors have power – usually backed by legislation – to oblige schools to act on their recommendations (Fig. 13.2).

In this figure, the opposites of *responsiveness* and *responsibility* reflect how, external evaluations call for a response by the school to an inspection report, while professional accountability focuses on the dimension of responsibility, particularly to the children in the reviewed schools. The degree of overlap in the white area concerns those cases where reviews sometimes mimic or become misappropriated as forms of rehearsal for an external review.

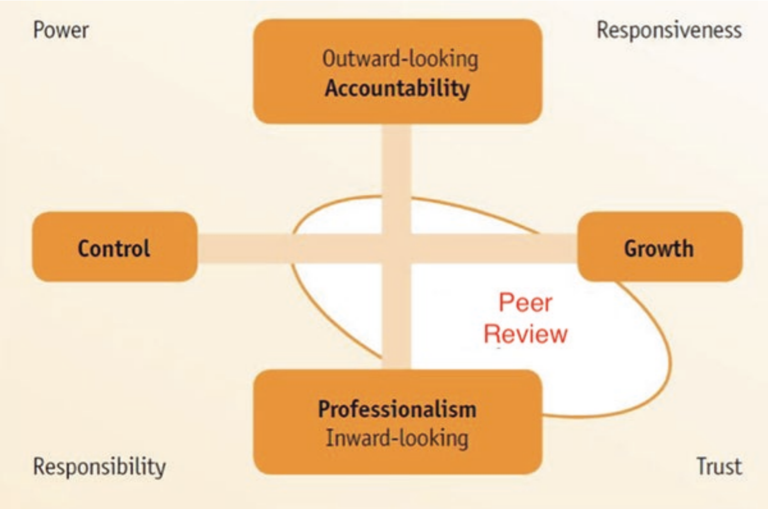


Fig. 13.2 A cognitive map of peer review. (Adapted from Gutknecht-Gmeiner 2007, p. 4)

13.4 Conclusions and Discussion

While there are commonalities between models, there are some important distinctions. The implications are multiple. For future research on peer review models, the ways in which the success of these programmes is measured may depend on the particular stance they take. For instance, building evaluation capacity may be the most important consideration in empowerment models while providing interim data on developmental targets identified in external evaluations may be more important in others, such as the Queensland model. For school leaders, the current capacity to evaluate will need to be taken into account; where this is low, schools may opt for models with more structured training and external facilitation. In schools with more capacity, building their own model of peer review with partner schools may be an option. Decisions on the focus of the peer review may determine which programme to adopt; schools may wish to share a themes or areas of pedagogy to develop or to set a wide scope in initial reviews, building familiarity and trust before engaging in more targeted enquiries.

Finally, for policy makers, the infrastructure and support required to integrate peer review systemically should be thought through. Finding alignment with existing external and self-evaluation practices is a balancing act; on the one hand alignment with existing evaluations may have benefits, while on the other hand, more subjective, constructivist approaches may be more contextually relevant and developmental.

Some further issues have emerged from the cases described in this book that are particularly worthy of further thought and analysis. In particular, the leadership development benefits of peer review, potential intermediate outcome of building collective efficacy among leaders and teachers, the mediating variable of trust, and other context factors, particularly in relation to research on networks.

13.4.1 Leadership Development Benefits of Peer Review

When speaking to school leaders involved in peer review it is not uncommon to hear them describe it as ‘the best professional development (PD) I’ve ever had’. This PD occurs incidentally, by virtue of gaining knowledge from other site visits and in professional discussion, through the direct observation of others’ practices and through the understanding accrued about what high quality education means from interviewing staff and students. Reviewers from outside of the host school can help to validate good practices, adopting an appreciative inquiry approach to build on this success and to provide a boost to staff morale. School leaders who conduct reviews develop their evaluation capacity, gaining the skills they need to evaluate initiatives brought into their own schools in the future.

Leadership development can be a deliberate intention of the peer review programme or the school leaders that opt to participate in reviews. In the case of the SPP model, selected middle leaders become improvement champions who design and facilitate a workshop to stimulate another recently reviewed school to implement improvements and to take ownership of these changes. The coaching model employed in this model provides excellent transferable leadership skills that can be used in their own site and as they move into more senior roles. In one of the Instead case studies (see Chap. 5), a headteacher used the peer review visit to prepare for her succession by handing over the task of receiving the review team and talking to them about the school to her deputy. Another Head wanted to place two newer middle leaders under pressure by practicing being held to account by other senior colleagues from another school.

Such professional development is consistent with Aubrey and Cohen’s (1995) organisational learning model of “working wisdom”. The model was stimulated by the answers the authors received from asking managers “Where did you learn what is most useful to you in working life?” Answers did not tend to involve training courses, but rather through involvement in a high-pressure project, or other career events. The authors suggested a number of principles of impactful PD, including a catalyzing event (usually being ‘nudged’ into one) and having numerous opportunities for dialogue among equals. Peer review programmes are particularly rich in both of these qualities.

A case can be made for peer reviewing as an emerging ‘signature pedagogy’ (Shulman 2005) for school leaders. A signature pedagogy has an intellectual, technical and ethical dimension (ibid). Those leaders participating in peer reviews often site the moral dimension of improving all children’s education, including those in partner schools. They also develop intellectual skills, through the justification of evaluative judgements based on careful scrutiny of evidence, and technical skills of coaching and giving feedback.

13.4.2 Collective Efficacy

Chapter 11 discusses in detail how the RiPR model of peer review has the potential to increase collective leader (CLE) and teacher efficacy (CTE). The same could be said of all of the models discussed in this book. CTE is said to be high when *school teams collectively feel confident that their own actions can, and do, have an impact on student learning*. High CTE is shown to be highly positively related to increased student attainment (Hattie 2016).

Without repeating the analysis here in detail, the logical case for examining the effects of peer review on CTE and CLE are strong. The professional dialogue, emphasis on collection of evidence and evaluation, and the ongoing support provided in peer reviews (especially when schools work in clusters or networks), is strong. Programmes such as SPP specifically try to increase collective responsibility, agency and efficacy as part of their theories of action. If peer review programmes succeed in increasing CTE in particular, then gains for children in their academic attainment are likely to follow. The challenge for peer review programmes centres on two key issues here: first, to ensure that evaluations penetrate to the level of teaching practices and second, to ensure follow up and support between schools. The extent to which programmes increase CTE is worthy of more attention as an outcome variable in future research on peer review programmes. The fact that most peer review programmes involve direct participation only of leadership teams means that while CLE may increase, this may or may not translate to increased CTE. The latter may be more dependent on the strategies adopted within each school, as prompted by the head teacher.

13.4.3 Trust as a Mediating Variable

The issue of trust has come out throughout the case studies in this book and indeed elsewhere in the literature on school collaborative work (e.g. Tschannen-Moran 2001). Trust can be seen as *the willingness to take a risk in a position of being vulnerable*, and to have five ‘faces’: *competence* (this has to do with professional level belief in others skills), *honesty* (match between deeds and statements), *openness* (sharing information), *benevolence* (presumption of good will) and *reliability* (the degree to which we can count on someone to come through with what is needed) (Hoy and Tschannen-Moran 1999). While all these conditions may be to a greater or lesser degree important, in peer review those of *openness* and *competence* appear most frequently mentioned. Without the willingness for all parties to share information about their schools openly, then the reviewers will struggle to reach adequate conclusions or to have the kind of professional dialogue that could stimulate school improvement. Additionally, schools being reviewed rely on the reviewing team to conduct their work thoroughly, fairly and accurately. If the reviewed school do not trust in the competence of the reviewers then they will clearly find it much harder to accept findings and recommendations of the review.

The need for reviewers to receive adequate training in evaluation is therefore important. In Chap. 2, Diamond and Kowalkiewicz comment:

> The credibility of the peer review was enhanced when there was strong reviewers' expertise and experience, as well as explicit reference to the NSIT (due to confidence in the tool). The potential for unreliable or invalid findings was reduced by involving reviewers with SIU experience or those known from existing professional networks, who were from outside the school and, on occasion, by increasing the size of the review teams (although this approach was not supported by all respondent principals) (Diamond and Kowalkiewicz, Chap. 2, p. xx).

The above extract suggests the potential advantages of using a standardized external evaluation framework that is known by all and to use reviewers with experience in this methodology. However, Greany's point in Chap. 4, about the process for schools becoming 'self-policing' rather than 'self-improving' needs to be taken into consideration. If reviewers adopt the approach of inspectors, this may help competency trust but damage other forms of trust, notably 'openness' where schools seek to prove to others that they are of a certain standard, choosing to withhold their weaknesses out of professional pride. Schools may mimic other schools that they see as excellent, due to their reputation gained through high inspection grade, leading to isomorphism rather than innovation. This would lead to a less developmental process in the review, inhibiting the scope for growth. The issue may be what the school hopes to gain from the review; Chap. 5 showed how some school leaders wanted to 'put staff on their mettle' through peer review in order to prepare them for the next inspection., while other schools may wish to develop areas not measured in external inspections. The question over the most appropriate lead reviewer will depend on the above factors and therefore this could be other credible school leaders in the network, credible school leaders outside the cluster, or specialist facilitators with expertise in evaluation and relevant research areas.

One of the questions about peer review networks concerns the degree to which schools can work in groups where there is not an *a priori* relationship of trust or whether there exists active 'mistrust'. Our early surveys on the SPP trial (Anders et al. n.d.) show how self-formed clusters come with a high level of trust between schools. This is not surprising with self-selected groups, and the programme organisers see this trust as essential pre-requisite to effective partnerships. However, establishing clusters that have not worked together or where one of the schools has greater needs than the others may lead to some very positive gains systemically. It is feasible that peer review relationships can help to build trust. Many of the examples given here show the positive reaction of participants in peer review schemes, so perhaps being eased into a partnership to experience its benefits could help establish enough trust to 'bare all' in a peer review visit.

The question of volition is interesting in peer reviews; in the English context some Multi-Academy Trusts (MATs) are mandating these for all schools in the trust; furthermore, in most peer review programmes the majority of the school staff, beyond senior leaders, will not have volunteered to take part in the programme. This again creates the question of whether obligatory peer review can work when trust may be low or if trust can be built if this is not pre-existing.

13.4.4 Network Configurations of Peer Review Programmes

Peer review programmes can occur within a variety of network level governance models and these may have an indirect effect on their effectiveness for school improvement.

Networks can be defined as "three or more legally autonomous organizations that work together to achieve not only their own goals but also a collective goal" (Provan and Kenis 2008, p. 231). These networks are 'goal-directed' rather than legally binding entities and thus for them to achieve outcomes above and beyond what they could as a collection of individual schools they would need to have some form of governance. Networks can have shared, *participant/shared governance*, where no separate entity is set up to oversee the network; a *lead organization-governed network* where one organisation, usually the most powerful is the lead in the network; or a *network administrative organization (NAO)*, where a separate entity oversees and facilitates the activities of the network.

Some may opt for a simple shared governance model, voluntarily grouping together with local partner schools to conduct periodic peer reviews and offer support. Others rely more on external facilitation by a peer review organisation or governmental agency to keep momentum and to broker school-to-school support. In the case of programmes such as CP and SPP, they form an NAO. Having a support organisation specifically designed to govern the network has the advantages of "enhancing network legitimacy, dealing with unique and complex network-level problems and issues, and reducing the complexity of shared governance" (Provan and Kenis 2008, p. 236). However, there is a degree of ambiguity here in that networks should have shared goals; although this may be true to a general extent with peer review organisations, the level at which schools are sharing goals is likely to be more at the individual cluster level (if at all). In the SPP model schools work in partnerships that adopt a looser participant governance role. These clusters of schools are encouraged to build in regular meetings to discuss the work of their network. Both the NAO and definitely the shared governance models of networking allow greater flexibility, can work well in small numbers of schools and benefit from a high degree of trust and goal consensus. However, the question arises whether this network can be seen to be the 'core business' of the schools or peripheral to its main activities.

Other network configurations include peer reviews that form part of the work of a formal multi-school organisation, for instance a MAT or Teaching School Alliance in England. Here, the trust organisation or Teaching School could be seen as the 'lead organisation' model of goverance (Provan and Kenis 2008, p. 235), that:

> provides administration for the network and/or facilitates the activities of member organisations in their efforts to achieve network goals, which may be closely aligned with the goals of the lead organisation (Provan and Kenis 2008, p. 235).

The benefits of the latter model are that schools in the network needn't have as much trust or goal consensus compared to the other two models. Where one of the schools in this network requires greater support, this could be more easily brokered and the potential to achieve network level outcomes and a sense of collective responsibility may be stronger in a more permanent network arrangement. Often (certainly in the case of England) schools are in hybrid arrangements with MATs or Teaching Schools Alliances (TSAs) using established peer review programmes with their schools, but trying to take ownership of the process, linking it in with their own network level evaluations and school improvement activities.[1]

13.4.5 Peer Review in a 'Coopetitive' Environment

Few countries have tried peer review as a system-wide strategy for school improvement. In Wales, Australia and particularly in England, nationwide growth in peer review use is most apparent. The ways in which these have grown to scale have much to do with the local educational environment. Some have described an environment in which intense competition is combined with equally strong incentives to collaborate. This so-called coopetition (Bengtsson and Kock 2000) leads organisations to find areas they feel comfortable in cooperating with partners in order to improve their effectiveness while avoiding working with other organisations on areas in which they may be competing directly.

In a report of local area partnership working in central England, some of the headteachers commented that "micro politics and local competition are perceived as too intense for high levels of trust and effective challenge to emerge" (Greany and Allen 2014, p. 34). This shows that even where peer review is becoming commonplace, the hierarchical and competitive environment may either affect the growth of peer review, or inhibit its potential for genuine school development. Greany, (Chap. 4) develops this point, observing that peer reviews have often been used as a form of 'self-policing' rather than 'self-improving. Some peer review projects have opted to work with schools just outside of their locality or with those in a different phase. A study in England (Stinton 2007), saw a peer review project that combined collaboration across the one FE college and three schools. Although all in the eastern region of England, they were sufficiently dispersed and this was said to help in that they were not competing for students and therefore able to be sufficiently open and honest in their appraisals.

---

[1] See case study at: https://audioboom.com/posts/6854494-matt-davis-and-marie-claire-bretherton-talk-school-to-school-collaboration

However these models evolve and adapt in different education systems, for schools to be truly self-evaluating and self-improving, peer review appears as a vital linchpin in the process. These collaborations are not sustainable without some form of infrastructure, continuous engagement, training and most likely, political will. While many countries may learn from the English, Welsh and Australian approaches to growing peer review, local adaptations will clearly be needed that reflect the different capacities of schools, network configurations and climate of trust. Peer reviews empower educational practitioners and encourage cooperation in what can be a keenly competitive field. The leadership development gained from leading and taking part in reviews of peer institutions is outstanding. The simple act of visiting another school to observe its practices is unusual for school leaders in many countries and is a formative experience for many. The deep professional dialogue about what is considered high quality education places the power to improve back into the hands of the profession rather than politicians or external evaluators.

Acknowledgements Many thanks to Dr. Rupert Higham who looked at an early draft of this chapter and provided some invaluable edits and feedback.

References

Argyris, C. (1976). Single-loop and double-loop models in research on decision making. *Administrative science quarterly*, 363–375.

Anders, J., Godfrey, D., Stoll, L., Greany, T., McGinity, R. & Munoz-Chereau, B. (n.d.). *Evaluation of the schools partnership programme*. Education Endowment Foundation.

Aubrey, R., & Cohen, P. M. (1995). *Working wisdom: Timeless skills and vanguard strategies for learning organizations*. San Francisco, CA: Jossey-Bass Publishers, Inc..

Bengtsson, M., & Kock, S. (2000). "Coopetition" in business networks—To cooperate and compete simultaneously. *Industrial Marketing Management, 29*(5), 411–426.

Campbell, D. T., & Stanley, J. C. (2015). *Experimental and quasi-experimental designs for research*. Ravenio Books.

Carr, W., & Kemmis, S. (2003). *Becoming critical: Education knowledge and action research*. Routledge.

Christie, C. A., & Alkin, M. C. (2012). Chapter 1: An evaluation theory tree. In M. C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (pp. 11–57). Sage Publications.

Cooperrider, D., & Srivastva, S. (2017). Appreciative inquiry in organizational life. In *Research in organizational change and development* (pp. 81–142). Emerald Publishing Limited.

Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis, 14*(4), 397–418.

Cronbach, L. J., & Shapiro, K. (1982). *Designing evaluations of educational and social programs*. Jossey-Bass.

Earley, P., & Weindling, D. (2004). *Understanding school leadership*. London: Sage.

Eisner, E. W. (1985). *The art of educational evaluation: A personal view*. Taylor & Francis.

Fetterman, D. M. (1994). Empowerment evaluation. Presidential address. *Evaluation Practice, 15*(1), 1–15.

Fetterman, D. M., Rodríguez-Campos, L., & Zukoski, A. P. (Eds.). (2018). *Collaborative, participatory, and empowerment evaluation: Stakeholder involvement approaches*. New York: Guilford Publications.

Greany, T., & Allen, T. (2014). *School improvement networks and system leadership in Coventry: Evaluating progress, areas for development and possible next steps*. London: London Centre for Leadership in Learning.

Gutknecht-Gmeiner, M. (2007). *Peer review in initial VET in Europe*. Report on the Pilot Phase and Outlook, Vienna.

Hattie, J. (2016). *Third annual visible learning conference (subtitled Mindframes and Maximizers)*. Washington, DC, July, 11, 2016.

Hoy, W. K., & Tschannen-Moran, M. (1999). Five faces of trust: An empirical confirmation in urban elementary schools. *Journal of School Leadership, 9*(3), 184–208.

Provan, K. G., & Kenis, P. (2008). Modes of network governance: Structure, management, and effectiveness. *Journal of public administration research and theory, 18*(2), 229–252.

Robinson, V. M., & Timperley, H. (2013). School improvement through theory engagement. In S. Kushner & M. Lei (Eds.), *A developmental and negotiated approach to school self-evaluation* (pp. 163–177). Emerald Group Publishing Limited.

Suggett, D. (2015). School autonomy: Necessary but not sufficient. *Evidence Base: A Journal of Evidence Reviews in Key Policy Areas, 2015*(1), 1–33.

Shulman, L. S. (2005). Signature pedagogies in the professions. *Daedalus, 134*(3), 52–59.

Stake, R. E. (1995). *The art of case study research*. Sage.

Stinton, J. (2007). *Evaluation of the college/school peer review pilot project*. Report No. RS7007, The Research Centre, CCN, Norwich.

Tschannen-Moran, M. (2001). Collaboration and the need fortrust. *Journal of Educational Administration, 39*(4), 308–331.

Tyler, R. W. (1942). General statement on evaluation. *The Journal of Educational Research, 35*(7), 492–501.

Dr. David Godfrey is an Associate Professor in Education, Leadership and Management at UCL Institute of Education in London and the programme leader for the MA Educational Leadership. He was co-director of the Centre for Educational Evaluation and Accountability until 2018 and was a lead inspector for the Independent Schools Inspectorate. An advocate of research-informed practice in education, his projects and publications include research-engaged schools, school peer review, inspection systems and lesson study. In July 2017, David was acknowledged in the Oxford Review of Education as one of the best new educational researchers in the UK.