

A review of practitioner-led evaluation of teacher professional development

Katrina McChesney* and Jill M. Aldridge

Curtin University

Western Australia

This is the authors' accepted manuscript. The Version of Record of this manuscript has been published and is available in *Professional Development in Education*, 45(2), 307-324.

<https://www.tandfonline.com/doi/full/10.1080/19415257.2018.1452782>

Sponsorship and Conflicts of Interest

The research reported in this article was supported by an Australian Government Research Training Programme Scholarship and an Australian Research Council Grant (LP110200285). The authors declare that they have no conflicts of interest.

A review of practitioner-led evaluation of teacher professional development

Abstract

The central role of teacher professional development in educational improvement is well-documented, and the investments of time and money into teacher professional development across the globe are immense. Given, however, that the quality of professional development has been shown to vary considerably, meaningful evaluation of professional development activities is critical and has been widely advocated by both researchers and education system leaders. Whereas past reviews have examined the nature and quality of researcher-led evaluations of professional development, the present review focuses on the nature and quality of practitioner-led evaluations. Using current theoretical recommendations for effective evaluation as benchmarks, the review examines the degree to which these recommendations are reflected in routine school-based evaluation of professional development, finding that a significant gap persists between theory and practice. The review also identifies practical and psychological factors that restrict further improvement in practitioner-led evaluation and proposes ways in which researchers, practitioners and policymakers could each contribute to future improvement in practitioner-led evaluation of professional development. Given that much teacher professional development occurs within school settings and/or is managed by school staff, improving the quality of routine practitioner-led *evaluation* of that professional development is a potentially powerful driver for positive change.

Keywords: Teacher professional development; Evaluation; Impact; Literature review; Theory and practice; Implementation

Introduction

Education systems around the world rely heavily on teacher professional development as a primary mechanism for achieving educational improvement (Darling-Hammond *et al.*, 2009; Bowe & Gore, 2016; Opfer, 2016). Extensive time and money are invested in teacher professional development (Hochberg & Desimone, 2010; TNTP, 2015; Cornman & Zhou, 2016), and numerous strategies are employed to promote teachers' participation in professional learning (Darling-Hammond *et al.*, 2009; Evans, 2014). However, despite the prominence of professional development as well as significant theoretical advances in identifying the types of professional development that are likely to improve teaching and learning (Borko, 2004; Timperley *et al.*, 2007; Opfer & Pedder, 2011; Hill *et al.*, 2013; Evans, 2014; Lauer *et al.*, 2014), the quality and impact of professional development offerings vary considerably (Darling-Hammond *et al.*, 2009; Opfer & Pedder, 2011; Hill *et al.*, 2013; Guskey, 2014; OECD, 2014; Ling & Mackenzie, 2015; TNTP, 2015; Bowe & Gore, 2016). These quality issues call into question the 'return on investment' associated with professional development and restrict the efficacy of professional development for generating the desired educational improvements. As such, it is imperative that effective evaluation is routinely carried out in order to facilitate improvements in both the quality and the outcomes of professional development (Guskey, 2000; Killion, 2003; Mitchem *et al.*, 2003; Earley & Porritt, 2010; Haslam, 2010).

There have been many calls for the routine practitioner-led evaluation of teacher professional development activities (see, for example, Guskey, 2000, 2002; Kelleher, 2003; Killion, 2003; Shaha *et al.*, 2004; Goodall *et al.*, 2005; Grossman & Hirsch, 2009; Earley & Porritt, 2014; Guskey, 2014; Porritt *et al.*, 2017). Further, numerous publications have explicitly attempted to equip school- and education system-based practitioners with knowledge and strategies to help them to conduct this evaluation effectively (see, for example, Australian

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

Institute for Teaching and School Leadership, n.d.; Guskey, 2000; Puma & Raphael, 2001; Killion, 2003; Speck & Knipe, 2005; Blank & de las Alas, 2008; Killion 2008; Bubb & Earley, 2010; Haslam, 2010; OECD, 2016; UK Department for Education, 2016). Given these efforts, the present review sought to examine the extent to which the calls and resources related to effective practitioner-led evaluation of professional development have been reflected in the routine evaluation of professional development by school-based practitioners.

It is important, at the outset of this article, to clarify our interpretations of the terms ‘professional development’ and ‘evaluation.’ In line with current trends, we took a broad view of what constitutes professional development: ‘Professional development is defined as activities that aim to develop an individual’s skills, knowledge, expertise and other characteristics as a teacher’ (OECD, 2014, p.86). Although alternative conceptualisations usefully describe professional development as change or growth within the practitioner (that is, the development of the professional; see, for example, Webster-Wright, 2009; Evans, 2011, 2014), a definition focused on professional development *activities* was appropriate for this review since it is such activities that commonly undergo evaluation. We note that the OECD (2014) definition provided above allows for the possibility that, although professional development activities may be *aimed at* developing professionals, they may not always be successful in achieving this development.

In terms of defining evaluation, we take the view that to ‘e-*valu*-ate’ something is to articulate its *value*—its worth, contribution or effects. Thus, in the context of professional development (and given the definition of professional development provided above), evaluation involves identifying and describing the value (worth, contribution or effects) of professional development activities. This stance necessitates a focus on *outcomes* or *impacts*—to evaluate is not merely to document that professional development happened or what it was like; rather, to evaluate is to articulate the outcomes or impacts of professional development.

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

As other scholars have previously noted, the ultimate goals of professional development are not the activities involved but the impacts that arise for teachers and learners (Guskey, 2000; Clarke & Hollingsworth, 2002; Guskey, 2002; Timperley *et al.*, 2007).

In considering the relationship between theory and practice in the area of teacher professional development evaluation, we note that the broader relationship between educational research and practice has long been a subject of debate and that, in recent years, much discourse has focused on improving this relationship (Kennedy, 1997; Gore & Gitlin, 2004; Biesta, 2007; Beycioglu *et al.*, 2010; Vanderlinde & van Braak, 2010; James, 2012; Jackson & Burch, 2016). There have been calls for: increased collaboration and dialogue between researchers and practitioners (Snow, 2015; Vanderlinde & van Braak, 2010); more deliberate efforts by researchers to listen (albeit critically) to the voices of teachers (Gore & Gitlin, 2004); further investigation of how research-practice partnerships can be most effective (Vanderlinde & van Braak, 2010; Coburn & Penuel, 2016); practice-oriented changes in our research priorities and paradigms (Sleeter, 2014; Bryk, 2015); and greater consideration of whether research outputs are useful or accessible for practitioners (Gore & Gitlin, 2004; James, 2012; Hill *et al.*, 2013; Bryk, 2015). It seems that, although ‘there is tension between production of knowledge for application versus for building a knowledge base, research in education is highly skewed toward the latter at the expense of the former’ (Sleeter, 2014, p.151). Moving beyond this theory–practice dichotomy, there is also a need for forms of practical theorising that bridge the theory–practice divide (Jackson & Burch, 2016). Given this backdrop, our hope, through this article, is to contribute to this collective reflection in terms of the specific case of practitioner-led evaluation of teacher professional development.

The Present Study

The overarching aim of this review was to examine the extent to which theoretical recommendations for evaluating teacher professional development have been implemented as

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

part of routine school-based professional development practice. Given this, three specific objectives were identified:

1. To provide a brief overview of current theoretical recommendations for best-practice evaluation of teacher professional development;
2. To review the available evidence regarding how teacher professional development is evaluated by practitioners and compare this with the theoretical recommendations; and
3. To identify barriers described in the literature that restrict the quality or extent of evaluation of teacher professional development in schools.

The next three sections of this article present the results corresponding to each of the above objectives. We conclude the article by discussing the implications of our findings and suggesting directions for future action.

In developing this article, we deliberately excluded analysis of the many professional development programme evaluations that are led by academics. Although many such studies are rigorous and reflect theoretical guidelines, the conditions that academics and school-based practitioners work under are very different. An important goal must be for school- and education system-based practitioners to implement high-quality evaluation practices; consequently, it was the evaluation approaches of practitioners that were considered to be relevant for inclusion in this review. For an excellent review of academic-led research in terms of professional development programme evaluation, we refer the reader to Hill *et al.* (2013).

Theoretical Recommendations for Evaluating Teacher Professional Development

Our first objective for this study was to provide a brief overview of current theoretical recommendations for best-practice evaluation of teacher professional development. In presenting such a review, we focus on *what* should be evaluated and *how* evaluation should be

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

conducted. Although these categories may appear to oversimplify a complex topic, they allowed us to target relevant literature-based guidelines that can be compared to what occurs in schools and education systems. The theoretical recommendations reviewed below may be applied to any form of professional development activity, ranging from traditional workshops and conferences to more modern approaches such as professional learning communities, teacher inquiry, social networking or teaching rounds.

What should be evaluated?

The core—and challenging—task of professional development evaluation has been defined as ‘translating the complex, interactive, formal, and informal nature of teacher learning opportunities into manageable, measurable phenomena’ (Desimone, 2009, p.183). To that end, a number of evaluation frameworks have been proposed by researchers (Guskey, 2000, 2002; Fishman *et al.*, 2003; Shaha *et al.*, 2004; Fraser *et al.*, 2007; Desimone, 2009; Bubb & Earley, 2010; Earley & Porritt, 2010; Coldwell & Simkins, 2011; Hunzicker, 2011; King, 2014); this section considers the components identified within these frameworks.

Across the various evaluation frameworks available, we identified three broad categories into which the component elements recommended for evaluation may be classified. First, some frameworks recommend evaluating *features of the professional development activity itself*. Objective features recommended for evaluation include the level of subject-specific content focus (Desimone, 2009); duration (Desimone, 2009; Earley & Porritt, 2010); opportunities for participant reflection and feedback (Earley & Porritt, 2010); whether the professional development was planned or incidental (Fraser *et al.*, 2007), one-off or ongoing (Hunzicker, 2011), or formal or informal (Fraser *et al.*, 2007); the links between theory and practice (Fraser *et al.*, 2007); and the degrees of collaboration (Desimone, 2009; Earley & Porritt, 2010; Hunzicker, 2011), active learning (Desimone, 2009), or instructional focus (Hunzicker, 2011). More subjective features of professional development activities that are

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

recommended for evaluation include the degree of professional relevance (Fraser *et al.*, 2007), coherence (Desimone, 2009), participant ownership (Earley & Porritt, 2010), intellectual stimulation (Fraser *et al.*, 2007), or supportiveness (Hunzicker, 2011) as well as the potential for teacher agency and transformative outcomes (Fraser *et al.*, 2007; King, 2014). Although any of these features may contribute to the impact of professional development, these features are not, in themselves, forms of impact. As such, we note that (based on our original definition of evaluation) examination of *only* these features would not constitute evaluation of professional development.

Second, some evaluation frameworks recommend examining the *contextual and strategic elements* associated with professional development. Such elements include baseline information (Bubb & Earley, 2010; King, 2014); details of the participants, school context or policy environment (Fraser *et al.*, 2007; Desimone, 2009; Coldwell & Simkins, 2011; King, 2014); the presence or nature of any goals, plans and targets (Fishman *et al.*, 2003; Bubb & Earley, 2010; King, 2014); the associated strategic leadership and vision (Earley & Porritt, 2010); and the degree and nature of organisational support for new practices (Guskey, 2000, 2002; Bubb & Earley, 2010; King, 2014). We see the same issue here as with the previous category—that is, although potentially informative and influential in terms of the impact of professional development, these contextual and strategic elements are not, in themselves, impacts and, therefore, cannot be the sole focus of evaluation efforts.

Finally, some (although not all) existing evaluation frameworks identify particular *impacts of professional development* that should be evaluated. The types of impact that are most widely identified for evaluation within the literature include teachers' affective reactions, teacher learning, teacher attitudes and beliefs, instructional changes and student outcomes (Guskey, 2000, 2002; Fishman *et al.*, 2003; Desimone, 2009; Bubb & Earley, 2010; Earley & Porritt, 2010; Coldwell & Simkins, 2011; King, 2014). Some evaluation frameworks also

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

incorporate additional types of impact including the cascading of new knowledge and practice to other staff and students (Bubb & Earley, 2010; King, 2014); changes in school processes or culture (Guskey, 2000, 2002; Coldwell & Simkins, 2011); and resource impacts such as time or financial costs (Shaha *et al.*, 2004). Of the three categories identified across the various existing evaluation frameworks, only the aspects reflected in this *impact* category address the requirements of our definition of evaluation.

Although we have emphasised the need for evaluation to be concerned with the outcomes or impacts of professional development, we acknowledge the value of information about the features of professional development and/or its context (the other two categories identified in our analysis). Such information can be of use in terms of understanding why particular professional development resulted in particular impacts as well as for planning future professional development activities. Many evaluation frameworks incorporate elements from at least two of the three categories identified above (see, for example, Guskey, 2000, 2002; Fraser *et al.*, 2007; Desimone, 2009; Bubb & Earley, 2010; Earley & Porritt, 2010; Coldwell & Simkins, 2011; King, 2014), and we do not doubt that measuring constructs across these categories may well produce a fuller picture of professional development than examining impacts in isolation.

It is of concern, however, that some frameworks (Fraser *et al.*, 2007; Hunzicker, 2011) omit consideration of any form of impact. Given our definition of evaluation, it is imperative that the actual impacts of professional development be examined; as such, frameworks not addressing impact need to be recognised as being insufficient (if used on their own) as a basis for evaluation. We do not deny that the frameworks by Fraser *et al.* (2007) and Hunzicker (2011) have a place: Fraser *et al.*'s (2007) lenses can usefully be employed to collect contextual information that can inform the interpretation of (impact-focused) evaluation data, and Hunzicker's (2011) checklist can be used to inform professional development planning.

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

However, in our view, it is important to clearly distinguish between *evaluation frameworks* (those that describe the evaluation of various forms of impact) and *non-evaluative frameworks* (those that serve different purposes related to the improvement and understanding of professional development but that do not evaluate impacts).

The justifications for evaluating many of the elements listed above (in all three categories) involve theoretical models of the teacher change process (see, for example, Supovitz & Turner, 2000; Guskey, 2002b; Desimone, 2009; Evans, 2011; Opfer & Pedder, 2011; Opfer, 2016; Boylan *et al.*, 2017). These models typically involve either a progressive (path), iterative (cyclical) or holistic picture of how various elements interact to generate teacher change. While detailing and comparing such models is beyond the scope of this article, we note that these models could be useful to professional development evaluators for both data analysis and interpretation, particularly in terms of examining where the teacher change process may have broken down and identifying areas that could be addressed to improve the outcomes of professional development. However, in using such models it is nonetheless important to distinguish between factors that influence, mediate or moderate the teacher change process (such as features of professional development opportunities and contextual and strategic elements) and those that reflect the impacts of that process.

How should evaluation be conducted?

In the context of an impact-focused understanding of evaluation, the question of how best to go about evaluating teacher professional development remains contested. Some researchers argue that experimental studies involving randomised controlled trials are the only way to definitively capture links between professional development activities and teacher or student outcomes (Yoon *et al.*, 2007; Whitehurst, 2012). However, another perspective calls for more flexibility in the face of the reality that, in most cases and contexts, such scientific approaches will not be practicable (Guskey, 2002, 2014). Thus, there are scholars who

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

advocate evaluation designs that, although empirical, systematic and reliable (to the extent possible), examine impacts in ways that are realistic and appropriate given the contexts and resources that are available (Guskey, 2002; Shaha et al., 2004; Desimone, 2009; Earley & Porritt, 2010; Bryk, 2015). For example, Guskey (2002, p.49) has recommended seeking ‘evidence, not proof’ of professional development’s impact. Within such approaches, however, it is important to ensure that the evidence collected is, nonetheless, substantiated and not merely asserted (Earley & Porritt, 2010).

The existing literature indicates that a range of data collection techniques can usefully contribute to evaluating the impacts of teacher professional development (Bredeson, 2002; Killion, 2008; Guskey, 2014). Teacher interviews, classroom observations and teacher surveys all have strengths and weaknesses and may—when designed and implemented appropriately—generate useful and reliable data (Desimone, 2009). Written assessments of teacher knowledge and student achievement data have become increasingly important and, although their use is not without controversy, such assessments have been perceived as objective measures that afford value-added analyses (Bryk, 2015; Hill *et al.*, 2013). Other techniques such as teacher reflections or portfolios afford deeper exploration of individual teachers’ professional learning (Guskey, 2002). Although, in principle, all of these techniques may have their place in professional development evaluation, only a limited number of specific instruments designed for this purpose have been reported over the past decade (Hill *et al.*, 2013), and not all of these have been made widely available for use by practitioners. As such, our review indicates that more work is needed to establish a corpus of thoroughly tested, widely-accepted and widely-available instruments.

At a broader level, the available literature highlights the importance of impact evaluation being embedded in the professional development ‘lifecycle’. There are widespread calls for evaluation to be planned for at the same time as the professional development

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

activities themselves are planned (Guskey, 2000; Mitchem *et al.*, 2003; Earley & Porritt, 2014) and for evaluation to specifically relate to the identified aims and goals of the professional development (Bredeson, 2002; Fishman *et al.*, 2003; Killion, 2003, 2008; Allison, 2013; Earley & Porritt, 2014). Finally, there are calls for the evaluation of professional development to be viewed formatively, rather than (solely) summatively, so that evaluation findings can inform the ongoing refinement of professional development practice, leading to increased levels of impact on teaching and learning (Killion, 2003; Mitchem *et al.*, 2003; Killion, 2008; Haslam, 2010; Guskey, 2014; Arbaugh *et al.*, 2016).

Evaluation practice

Our second research objective was to review the available evidence regarding how teacher professional development is evaluated ‘on the ground’ by practitioners and to compare this with the theoretical recommendations reviewed above. The results of our analysis are described below in terms of *what* is evaluated and *how* evaluation is conducted—the same categories that were used to organise the review of theoretical recommendations. It is important to note that, although we searched for literature from any geographic context and had hoped to provide a comprehensive international review, the majority of the literature located was UK-based.

What gets evaluated?

Our review of literature indicated that the factors considered in many school-based evaluations of professional development are in contrast to the literature-based recommendations outlined above and to our definition of evaluation as being focused on the impact or outcomes of professional development. School-based ‘evaluations’ of professional development appear to frequently exclude examination of critical impacts such as teacher learning, changes in classroom practice and student outcomes. This may not be surprising to

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

many readers; however, it is of concern that evaluation practice has not progressed significantly despite the widespread calls for improvement over the past two decades (see Introduction).

Perhaps the largest and most detailed examination of school-based evaluation practices to date, carried out in the UK by Goodall *et al.* (2005; see also Harris *et al.*, 2006; Muijs & Lindsay, 2008), found that participant satisfaction with professional development was *usually* or *always* evaluated in at least 75% of the 1000 schools involved. Other elements (including impacts on teaching and learning) were evaluated far less frequently: ‘According to teachers, in the average school only participant satisfaction is likely to be evaluated’ (Muijs & Lindsay, 2008, p.205). Although participant reactions to professional development can be conceptualised as impacts of professional development (see, for example, Guskey, 2000, 2002; Bubb & Earley, 2010; King, 2014), measuring only this type of impact provides an incomplete picture that excludes core teaching and learning impacts.

Although Goodall *et al.*’s (2005) study is now over 10 years old, several subsequent UK-based studies have reached similar conclusions. For example, drawing on a series of UK school inspections, Ofsted (2006) noted that effective evaluation of professional development occurred in only 13 of the 29 schools inspected. Four years later, based on a follow-up study of 40 schools specifically selected for their good practice in the area of professional development, Ofsted once again found that, even in those high-performing schools, the evaluation of professional development (in terms of its impact and its value for money) remained ‘the weakest aspect of continuing professional development’ and needed significant improvement (Ofsted, 2010, p.5). Likewise, the 2008 *State of the Nation* study of teacher professional development in England found that evaluation remained limited in scope, typically focusing only on teacher reactions and neglecting teaching and learning impacts (McCormick, 2010; Pedder & Opfer, 2010). In a review of the work of professional development providers in the

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

UK during 2010-2011, over 99% of cases reviewed reported some ‘evaluation’ of professional development activities (CUREE, n.d.). However, this ‘evaluation’ was frequently focused on the format and content of professional development activities, whereas

in a large minority of provision (27%), there was nothing in place which meant participants would be equipped to assess the impact of their professional development in terms of their students’ learning. (CUREE, n.d.).

Most recently, case studies conducted in relation to the UK *Effective Practices in Continuing Professional Development* project confirmed that many schools still ‘struggle[d] with how to demonstrate and evidence the impact of PD [professional development]’ (Earley & Porritt, 2014, p.116) and, therefore, focused their evaluations on the processes and products (such as policies and resources), rather than the impacts, of professional development. It was found that ‘for many of the project participants, linking PD activity to desired pupil outcomes was not yet a natural or expected practice’ (p.118).

Only one UK-based study contradicts the above findings. In a study involving over 1,500 school-based professional development leaders, ‘most respondents (96 per cent) indicated that their school either “always” or “sometimes” evaluated the impact of CPD’ (Robinson *et al.*, 2008). While this appears to be a promising finding, it should be noted that this study (unlike those reported above) relied on professional development leaders’ self-reports of their evaluation practices.

Outside the UK, in addition to evaluation of teachers’ affective reactions to professional development (Grossman & Hirsch, 2009), much ‘evaluation’ practice appears to centre on examining either *participation* in professional development or the *features* of professional development activities. A focus on teachers’ *participation* in professional development is

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

evidenced in both Australia (Ling & Mackenzie, 2015) and the US (Grossman & Hirsch, 2009; Barrett *et al.*, 2013). For example, Barrett *et al.* (2013, p.351) report that most US states:

requir[e] teachers to participate in some form of professional development every year as part of their continuing certification or licensure requirements. We are unaware of any state that requires programs to prove their effectiveness in order to be included in the list of acceptable professional development options ... and no state has incorporated mechanisms that match teachers to professional development programs based on improving teacher effectiveness.

In such a context, evaluating only teachers' participation in professional development without considering the nature (features), the relevance, or—critically—the impact of that professional development raises serious cause for concern.

On the other hand, a focus on *features* of professional development is evident in the US (Birman *et al.*, 2009), where professional development initiatives are primarily evaluated using teachers' descriptions of the features of the professional development that they have experienced, such as its instructional focus, degree of active learning, or duration. Like evaluating only teachers' participation in professional development, evaluating only the features of professional development is an approach that reflects portions of the literature-based guidance (see, for example, Garet *et al.*, 2001; Desimone, 2009) yet ignores other guidance that emphasises the importance of also evaluating professional development's impact (Guskey, 2000; Earley & Porritt, 2014; Guskey, 2014). Crucially, neither of these approaches meets our impact-focused definition of evaluation (see Introduction).

Although it is concerning that large numbers of schools are not evaluating the impacts of professional development, pockets of excellence can nonetheless be identified. For example,

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

Ofsted's (2006, 2010) reports identified examples of good practice, noting that the schools that evaluated most effectively went beyond superficial measures of implementation compliance (such as the frequency of teachers' use of a particular teaching technique) and considered the holistic impacts of professional development on teaching, learning and school life.

A further positive finding was that although schools may not have been consistently evaluating key aspects of professional development, there did at least appear to be a growing awareness of the need to do so. For example, Earley & Porritt (2014) found that although schools did not consider the effects of professional development on students, there was at least some consideration of the teacher learning that might result from professional development. Further, two examinations of evaluation practices in the US suggested that states and school districts recognised the need to consider the teaching and learning impacts of professional development and that although they were not carried out, some schools and districts had developed sound plans for evaluating these impacts (Blank *et al.*, 2008; Regional Education Laboratory Southeast, 2009).

How is evaluation conducted?

According to the literature reviewed, by far the most common school-based approach to evaluating professional development remains the use of an evaluation form to examine teachers' immediate reactions (usually affective reactions) at the end of a professional development activity (Goodall *et al.*, 2005; Robinson *et al.*, 2008; Regional Education Laboratory Southeast, 2009; Pedder & Opfer, 2010; Earley & Porritt, 2014). There was no evidence to suggest a decline in the use of evaluation forms over time. Participant satisfaction can be understood as an (albeit preliminary) form of impact and readers are unlikely to be surprised that such evaluation forms remain widely used. However, it is widely recognised that this type of evaluation is likely to generate only a limited body of highly subjective data and offers no information about the impact of professional development activities (Guskey, 2000).

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

As such, it is of concern that, in many cases, satisfaction-focused evaluation forms appear to be the *sole* evaluation mechanism used (Goodall *et al.*, 2005; Robinson *et al.*, 2008; Regional Education Laboratory Southeast, 2009; Pedder & Opfer, 2010; Earley & Porritt, 2014).

Besides evaluation forms, there was evidence that some schools merely documented (as their ‘evaluation’) practical details about professional development activities, such as the courses that were run and their dates, times and attendees (Guskey, 2000; Killion, 2008; Regional Education Laboratory Southeast, 2009). Reliance on this technique omits direct examination of impact and instead suggests an assumption that teachers’ participation in professional development or covering relevant topics will naturally generate positive outcomes (Killion, 2008). There was also evidence that some schools tracked the dissemination of information and materials from professional development activities and reported this tracking as constituting evaluation (Brown *et al.*, 2001; Goodall *et al.*, 2005; Harris *et al.*, 2006). Although there may be value in staff, upon returning from off-site professional development activities, sharing what they have learned with colleagues, documenting this process does not provide information about either the quality of the professional development activity or its impact on teaching and learning.

The evidence indicating that schools used either the reporting of practical details or the dissemination of professional development content as proxies for evaluation dates only to 2009. We did not find evidence of these strategies being used more recently, although this does not necessarily mean that they are not still in use. Further investigation into the persistence of these two ‘evaluation’ approaches is, therefore, necessary.

The literature reviewed also indicates that some school leaders rely on unsupported assertions or anecdotal evidence to determine the impact of professional development (Ofsted, 2010; Earley & Porritt, 2014). For example, Earley & Porritt (2014, p.116) reported that, across 20 case study schools, most professional development leaders ‘were able to assert that

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

their project had made a difference but were unable to substantiate that or explain how the PD in which participants had engaged had brought about demonstrable improvements in classroom practice and student learning'. Data sources such as lesson observations or interviews with staff were sometimes named as evidence of impact, yet details of what these sources actually showed were consistently lacking. Ofsted (2010, p.27) further noted that such an approach 'sometimes led to a more positive view [of professional development's impact] than was warranted.'

In both the US and the UK, there is evidence of some professional development impact evaluation being undertaken in the context of researcher–practitioner partnerships (Blank *et al.*, 2008; Casteel & Ballantyne, 2010; Earley & Porritt, 2010). Although the focus of our review was practitioner-led evaluation, researcher–practitioner partnerships are worth noting here because of their potential influence on practitioner-led evaluation. The literature indicates that such researcher–practitioner partnerships vary in their success. On the one hand, Blank *et al.* (2008, p.2) report that 'Our analysis of evaluation findings across a number of programs and studies indicates that partnerships between higher education institutions and school districts have generally not added to the capacity for evaluation of professional development'. On the other hand, Earley & Porritt (2010, p.142) document a set of in-depth case studies of research–practice partnerships across which there is evidence that:

Many [schools] made significant progress ... especially in understanding that impact evaluation is not an end in itself but a way to improve the quality of the intended outcomes. Project leaders were able to make use of a greater variety of methods to gather impact data, to identify what they were learning from that evidence and demonstrate greater clarity about the improvements they wanted to achieve.

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

It appears, then, that where researcher–practitioner partnerships work well, they have the potential to improve both evaluation practices and the knowledge and skills of school-based practitioners.

Overall, based on the evidence available, the approaches used by practitioners for evaluating teacher professional development appear to be limited in both range and scope. With the exception of (some) researcher–practitioner partnerships, typical evaluation practice seems not to reflect the theoretical recommendations for either *what* should be evaluated or *how* evaluation should be conducted. As such, school-based professional development evaluation appears still to be ‘the weakest link in the [professional development] chain’ (Ofsted, 2006, p.19). This is of concern since without accurately measuring the quality or impact of teacher professional development, practice is unlikely to improve.

Barriers to effective evaluation

Our final research objective was to identify barriers described in the literature that restrict the quality or extent of evaluation of teacher professional development in schools. The evidence reviewed above reveals a significant discrepancy between theory and practice in this area, yet theoretical recommendations for evaluating teacher professional development have been established for some time, and many supporting resources are available to schools (see, for example, Australian Institute for Teaching and School Leadership, n.d.; Guskey, 2000, 2002; Shaha et al., 2004; Speck & Knipe, 2005; Blank & de las Alas, 2008; Killion, 2008; Bubb & Earley, 2010; Haslam, 2010; Guskey, 2014). It is important, therefore, to investigate why the discrepancy between the theory and practice of professional development evaluation persists.

Based on our review of literature, we identified both practical and psychological barriers that may prevent the widespread implementation of effective evaluation practices. We

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

review these barriers below. It should be noted that, in our opinion, these barriers are challenges that need to be overcome, rather than justifications for accepting the existing disparity between theory and practice.

Practical barriers

Perhaps the most obvious practical barrier to the high-quality evaluation of teacher professional development is the time, effort and cost required to enact best-practice recommendations (TDA, 2007; Robinson *et al.*, 2008; Desimone, 2009). It is unlikely that the time-consuming and costly techniques generally used in rigorous, academic-led evaluations of teacher professional development (such as interviews and classroom observations) will be realistic options for widespread use in schools (Blank *et al.*, 2008). Further, this high resource demand is multiplied when we consider the plethora of professional development, both formal and informal, that occurs in schools—all of which should be evaluated in terms of its impact (Guskey, 2000).

A second practical barrier to effective evaluation practice is the complex nature of school environments (Guskey, 2002; Bryk, 2015). Within schools, there are many factors such as improvement initiatives, contextual influences and ‘ambient professional development’ (Wayne *et al.*, 2008, p.472) that interact, overlap and affect each other, making it difficult to determine whether observed impacts were the result of particular professional development.

Given the factors described above, practitioners require impact evaluation tools that are ready-to-use, theoretically sound, as well as manageable and appropriate for use. Schools have reported a lack of such tools (Borko, 2004; Goodall *et al.*, 2005; Blank *et al.*, 2008; Desimone, 2009), and, indeed, some of the new evaluation tools published in the academic literature are likely either to be beyond the scope of what is manageable for routine school-based use (see, for example, Trygstad *et al.*, 2014) or else to provide aggregate data that is not informative for

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

evaluating or improving individual professional development activities (see, for example, Soine & Lumpe, 2014). We have argued elsewhere (McChesney & Aldridge, 2018) that it is important to develop tools that suit the needs of practitioners, and we suggest that there is a place for ‘generic’ tools that can be used to evaluate a range of forms of professional development.

Finally, there is some evidence of capacity issues among school personnel who are tasked with leading and evaluating professional development (TDA, 2007; Robinson et al., 2008; Ofsted, 2010; OECD, 2016). These issues relate to professional development leaders’ capacity for both impact evaluation in particular and the collection, analysis and use of evidence and data to inform practice in general (OECD, 2016).

Psychological barriers

Psychological barriers may also inhibit the effective evaluation of teacher professional development. Some such barriers emerge as teachers and school leaders seek to avoid positions of vulnerability. For example, teachers who have been required to attend professional development (non-volunteers) may not be willing to cooperate with evaluation techniques (such as written assessments of their learning) due to fear or suspicion about how the results will be used (Fishman *et al.*, 2003; Le Fevre, 2014). The extensive criticism that teachers, collectively and individually, receive because of the perceived failings of education may also influence teachers’ willingness (or lack thereof) to participate in investigations of how professional development impacts on themselves and on their students (Wilson & Berne, 1999; Desimone, 2009).

In the current accountability context of education, negative findings regarding the impact of professional development efforts could pose a risk to schools and education systems that may be competing for funding, ratings, enrolments, or other factors (Shaha *et al.*, 2004;

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

Hochberg & Desimone, 2010). This, coupled with the knowledge that, at present, much professional development does not have a sizeable impact on teaching and learning, may lead school administrators to favour the ‘safer’ (but less rigorous and, ultimately, non-evaluative) methods described earlier. Although, ideally, schools should value accurate formative feedback to inform their practice, Wilson & Berne (1999) suggest that we should not be surprised that schools or systems choose to simply report on the activities and processes of professional development, as ‘[p]rocess is the single thing that one can count on’ (p.199).

Discussion

Our review has shown that theoretical recommendations for evaluating teacher professional development are, in many cases, not reflected in the routine school-based evaluation of professional development. In highlighting concerns regarding practitioner-led professional development evaluation, we do not discount the many rigorous evaluation studies that have been conducted by educational researchers (see, for example, Hill *et al.*, 2013). These studies can, and do, benefit both the research community and school-based practitioners (particularly in cultural terms; see Biesta, 2007). Our interest, however, was to go beyond these researcher-led studies to consider the evaluation that is occurring in the majority of schools and education systems—evaluation that is led by practitioners and restricted by a different set of complexities and constraints. This larger, school- and system-based body of evaluation, if done well, has the potential to facilitate substantial and widespread improvement in professional development practice, resulting in greater impacts on educational processes and outcomes. The return on investment, then, makes such practitioner-led evaluation worth supporting.

In terms of the theoretical base for this field, we note that although they have been established for some time, the existing theoretical recommendations for professional development evaluation are neither perfect nor comprehensive. Constructivist views of the individualised nature of learning indicate that the nuances of teacher learning and the

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

pathways, affordances, barriers and moderators affecting the outcomes of teacher professional development can never be accurately reduced to a set of ‘manageable, measurable phenomena’ (Desimone, 2009, p.183). Within the research community that is associated with the evaluation of teacher professional development, disagreement and mutual critique persist, particularly around methodological considerations such as the importance of experimental or quasi-experimental approaches in evaluating teacher professional development. Nonetheless, even as theoretical recommendations continue to be refined over time, it is desirable that practitioners increasingly engage with and implement the best available guidance.

Our review sits within an international context of somewhat increasing policy attention to the impacts of professional development. The OECD has attempted to draw governments’ attention to the quality and impacts of professional development (OECD, 2009; Barrera-Pedemonte, 2016; Opfer, 2016), and several education systems have established policy frameworks that address impact evaluation for teacher professional development (see, for example, General Teaching Council for Scotland, n.d.; Australian Institute for Teaching and School Leadership, 2012; The Teaching Council, 2016). On the other hand, many education systems appear to remain focused on promoting teacher *participation* in professional development (or in particular types of professional development that are believed to be more effective) with minimal attention to impact evaluation (New Zealand Ministry of Education, n.d.; Grossman & Hirsch, 2009; Barrett *et al.*, 2013; Ling & Mackenzie, 2015).

Unfortunately, our review indicates that, as yet, there has not been substantial improvement in either the scope (what is evaluated) or approaches (how evaluation is conducted) used in school-based evaluation practice since the turn of the century, despite the widespread calls for such improvement and the availability of numerous publications aimed at supporting school-based evaluation practice (see Introduction). Some instances of good evaluation practice have been identified (Ofsted, 2006, 2010), and initiatives such as Earley &

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

Porritt's (2014) *Effective Practices in Continuing Professional Development* project have successfully raised school capacity in the area of professional development impact evaluation while also reporting usefully on the current state of practice in this area. What is still needed, however, is for such improvement in evaluation practice to occur on a sufficiently large scale that it becomes normative; achieving this is likely to require explicit efforts to address the barriers identified in this review.

The literature we reviewed indicated that one key barrier for practitioners is the lack of practicable evaluation tools. Although this is a consistent message from practitioners, we nonetheless located at least *some* impact evaluation tools or frameworks that, in our opinion, do appear to be practicable for routine use in school contexts (see, for example, Mullins *et al.*, 2010; TDA, 2010; see also McChesney & Aldridge, 2018 and the overview provided in TDA, 2007). Further, we located auditing or accreditation initiatives—such as the UK-based *Investors in People*¹ programme and the Teacher Development Trust's *CPD Audit*²—that offer further practical support and guidance for schools in the area of professional development evaluation. As such, it is possible that the key issue may be less about the *availability* of appropriate resources and more about practitioners' *access* to or *awareness* of them. Indeed, Robinson *et al.* (2007) reported that school-based professional development leaders relied primarily on in-school information and resources to inform their practice, to a lesser extent on local and regional sources, and to an even lesser extent on national-level provision, suggesting 'closed loop' information systems that do not afford access to the full range of available resources.

A limitation of this review relates to our finding that the greatest amount of research attention to practitioner-led evaluation practices for teacher professional development has been

¹ See <https://www.investorsinpeople.com/resources/share-and-inspire/investors-people-schools-academies-and-colleges>

² See <http://tdtrust.org/evaluation>

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

focused in the UK. Far less information is available for other contexts, making our inferences outside the UK context less conclusive. As such, there is a need to expand, and add greater detail to, the literature base describing school-based evaluation of teacher professional development in a range of international contexts.

Future directions

Given our findings regarding the current state of theory and practice in the area of teacher professional development evaluation, we now offer some initial thoughts on possible ways forward. In doing so, we acknowledge that many factors can contribute to disconnection between educational research and practice, including both practitioners' and researchers' attitudes as well as the socio-political landscape and features of the educational system itself (Kennedy, 1997; Vanderlinde & van Braak, 2010). As such, increasing the alignment between theory and practice in this specific area will be a complex and challenging undertaking. We do not imagine that the suggestions below are either definitive or comprehensive; rather, we hope that they will facilitate discussion around how the various stakeholders (researchers, practitioners and policy-makers) might contribute to improving the evaluation of teacher professional development in practice.

To avoid perceptions of academic superiority or a desire to 'place blame', we first consider how the research community—to which we, ourselves, belong—may be able to contribute: What might be within our own domain of influence that could support improvement in evaluation practice? Part of our work here may be philosophical: Perhaps, in addition to the existing discourse around the importance and refinement of optimal, comprehensive, rigorous evaluation approaches and methods—*What is best?*—a new conversation might emerge, exploring: *What evaluation approaches and methods remain valid, useful and theoretically-informed yet are genuinely practical and manageable for routine, widespread use by practitioners in real school contexts?* We imagine that there will be different perspectives on

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

the ideal balance between theoretical and practical considerations (or on whether this is actually a trade-off at all) and suggest that this warrants deep consideration and discussion.

A related discourse that could develop might further probe the existing calls for routine evaluation of professional development by practitioners, asking: *What should we realistically expect from schools' routine evaluation of professional development, why, and what support would be needed for this to occur?* An extreme view was stated by Flecknoe (2002) who suggested that, in light of the difficulties involved, we should let go of the 'unrealistic' (p.119) expectation that schools and education systems evaluate the impact of professional development. How do we, as educational researchers, respond to such suggestions? The calls by many researchers, including Guskey (2000, 2002, 2014), Kelleher (2003), Shaha *et al.* (2004), Goodall *et al.* (2005), and Earley & Porritt (2014), indicate that meaningful evaluation of the impact of most, if not all, professional development practice is critical for maximising the contribution of professional development to educational improvement goals. Therefore, rather than giving up on practical evaluation, we should seek new ways to make routine evaluation increasingly manageable and robust. Work that draws on the field of evaluation capacity-building (part of the business field of organisational management) and considers how this might transfer to educational contexts may represent one fruitful line of inquiry (King, 2002; Lawrenz *et al.*, 2011; Sumsion *et al.*, 2015).

From a more practical standpoint, educational researchers might usefully seek to develop, validate, refine and disseminate new impact evaluation tools that balance theoretical principles with practical constraints. Such work could be informed by the new strands of discourse proposed above and would help to fill a gap reported by schools (Borko, 2004; Goodall *et al.*, 2005; Blank *et al.*, 2008; Desimone, 2009). Careful consideration of how to ensure practitioners are aware of any new tools that may be developed will also be important.

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

In terms of developing new tools, although Earley & Porritt (2010) have noted the prevalence of unsubstantiated assertions about the impact of professional development, we share Desimone's (2009) view that teacher self-report data remains a valid way of examining the impact of professional development (see also McChesney & Aldridge, 2018). Such self-report data can include examples or descriptions of specific teacher learning, changes in classroom practices, or changes in student outcomes, addressing Earley & Porritt's (2010) concern about asserted versus substantiated evidence. Given the importance of teacher agency and buy-in for effective professional development and change (Kennedy, 2005, 2014; McChesney, 2017), evaluation tools and practices that respect teachers' professionalism and invite them to *give an account* (rather than insisting that they be *held to account*; see OECD, 2016; Lingard *et al.*, 2017) seem most likely to be effective.

Finally, as researchers, we might also shift somewhat in our consideration of both research topics and research designs, giving greater priority to the needs of practitioners. Hill *et al.* (2013) have already proposed a new research model involving small-scale, rigorous initial trials to refine the design of a professional development initiative. Within this model, existing empirical, positivist/post-positivist research methods are leveraged in new ways to provide more practically useful information at an earlier stage. More broadly, there have long been calls for the identification of educational research priorities to be based on practical needs rather than gaps in theoretical literature (Kennedy, 1997; James, 2012; Sleeter, 2014; Snow, 2015, 2016). The evaluation—and, ultimately, the quality and impact—of teacher professional development by practitioners could be significantly improved if such shifts were to occur.

Moving beyond what educational researchers alone could contribute, our review indicated that partnerships between researchers and practitioners may also facilitate meaningful improvement in practitioners' evaluation of teacher professional development. Although some previous partnership approaches to professional development evaluation have not been

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

effective for increasing practitioners' evaluation capacity (Blank *et al.*, 2008) and much is not yet known about how to make such partnerships productive (Coburn & Penuel, 2016), we are nonetheless encouraged by the positive reports of researcher–practitioner partnerships provided by Casteel & Ballantyne (2010) and by Earley & Porritt (2014). Looking forward, such partnerships may be potential 'spaces of freedom' where better ways of working can, and should, be pursued (Gore & Gitlin, 2004, p.53). Researcher–practitioner engagement may also be an effective forum for raising practitioners' buy-in around the importance and value of improved evaluation practice.

Practitioners working in schools and engaging in professional development evaluation also have a part to play in improving practice. If the first step in any change process is knowledge (Rogers, 1995), then school-based evaluators of professional development need knowledge in three areas. First, practitioners need to become aware that there are issues with the evaluation approaches that they may currently be using. Second, practitioners need to become more familiar with the theoretical recommendations for evaluating professional development—particularly in terms of the importance of directly evaluating impact. Third, practitioners need to be aware of the evaluation resources that are currently available. With such a foundation, practitioners could begin to critique and experiment with evaluation approaches, reflecting on what might be theoretically-sound yet manageable and meaningful in their school contexts. Practitioners could then function as researchers themselves, inquiring into their own practice as they seek to find ways of effectively evaluating professional development and then sharing their findings.

A final group of stakeholders who could contribute to addressing the disparity between the theory and practice of professional development evaluation comprises policy-makers and education system leaders. We would, of course, caution against reactionary policies demanding intensive, comprehensive evaluations of all professional development efforts using methods

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

that, while rigorous, could place unrealistic burdens on schools and education systems. However, our review indicates that perhaps the most progress in improving routine school-based evaluation of professional development has occurred in the UK, where managerial modes of accountability (Lingard *et al.*, 2017) and the associated policies for teacher professional development, performance evaluation, school self-evaluation and workforce management have driven changes at teacher, school and system levels. As such, considered policy changes in other contexts may facilitate improvement in practitioners' awareness and practice around evaluating teacher professional development; as the old adage says, 'what gets measured gets done'.

Policy-makers might also advance the field by commissioning their own projects, or funding proposed studies, aimed either at investigating practicable yet theoretically-informed evaluation models or at identifying ways to make research-practice partnerships more effective. Policy-makers might also actively seek to support increasing awareness among practitioners regarding the need for improvement in evaluation practice, existing theoretical recommendations, and available evaluation resources. Finally, with time, as new and appropriate strategies for routine evaluation of the impact of professional development emerge, these may be able to inform or be embedded in policy, in order to increase their dissemination and implementation.

Conclusion

In this review, we have examined the state of practitioner-led evaluation of teacher professional development. To our knowledge, no similar review has previously been published, despite the well-documented reliance on professional development for improving teaching and learning in education systems across the globe.

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

In conducting our review, we have identified that there appears to be a significant disconnect between the existing theoretical recommendations and the ‘evaluation’ approaches commonly used in schools. This is an important finding, as it highlights an area that urgently needs addressing—and one that, if addressed, could lead to improvement in both the quality and the impact of teacher professional development. To facilitate this process, we have offered a number of suggestions regarding how researchers, practitioners and policy makers might each contribute to this improvement.

An important contribution of our review is our examination of specific barriers that, to date, have contributed to the gap between theory and practice in this area. It is only through recognising these barriers that we can begin to seek solutions and strategies that will not continue to be undermined by those very barriers.

We acknowledge that, as stated by Snow (2016, p.67), ‘starting from urgent problems of practice and working in partnership with practitioners, one encounters and generates many obstacles’. Nonetheless, it is clear from our review that there is a need for improvement in the area of teacher professional development evaluation. We have identified that there are serious issues with the ways in which evaluation is typically conducted, as well as barriers that inhibit the implementation of existing theoretical guidance. In our view, this falls within Schoenfeld’s (1999, p.5) classification of ‘problems whose solutions help make things better and contribute to theoretical understanding’ and, as such, working towards addressing this issue may be ‘a high-leverage strategy for making a difference in the years to come.’

REFERENCES

- Allison, A. (2013). No art teacher left behind: Professional development that really matters in an age of accountability. *Arts Education Policy Review, 114*(4), 178-190.
- Arbaugh, F., Marra, R., Lannin, J.K., Cheng, Y., Merle-Johnson, D., & Smith, R. (2016). Supporting university content specialists in providing effective professional development: The educative role of evaluation. *Teacher Development, 20*(4), 538-556.
- Australian Institute for Teaching and School Leadership. (n.d.). *The essential guide to professional learning: Evaluation*. Retrieved from <https://www.aitsl.edu.au/tools-resources/resource/the-essential-guide-to-professional-learning-evaluation>
- Australian Institute for Teaching and School Leadership. (2012). *Australian charter for the professional learning of teachers and school leaders*. Carlton South, Australia: Author.
- Barrera-Pedemonte, F. (2016). *High-quality teacher professional development and classroom teaching practices: Evidence from TALIS 2013*. Paris, France: OECD.
- Barrett, N., Butler, J. S., & Toma, E. F. (2013). Do less effective teachers choose professional development does it matter [sic]? *Evaluation Review, 36*(5), 346-374.
- Beycioglu, K., Ozer, N., & Ugurlu, C.T. (2010). Teachers' views on educational research. *Teaching and Teacher Education, 26*(4), 1088-1093.
- Biesta, G.J.J. (2007). Bridging the gap between educational research and educational practice: The need for critical distance. *Educational Research and Evaluation, 13*(3), 295-301.
- Birman, B.F., Boyle, A., Le Floch, K.C., Elledge, A., Holtzman, D., Song, M., . . . Yoon, K.S. (2009). *State and local implementation of the No Child Left Behind Act, Volume VII - Teacher quality under NCLB: Final report*. Washington, DC: US Department of Education.
- Blank, R.K., & de las Alas, N. (2008). *Current models for evaluating effectiveness of teacher professional development: Recommendations to state leaders from leading experts*

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

- (*Summary report of a CCSSO conference*). Retrieved from <http://programs.ccsso.org/content/pdfs/Current%20Models%20for%20Eval%20Effect%20of%20Teacher%20PD%20summary%20Report.pdf>
- Blank, R.K., de las Alas, N., & Smith, C. (2008). *Does teacher professional development have effects on teaching and learning?* Washington, DC: Council of Chief State School Officers.
- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, 33(8), 3-15.
- Bowe, J., & Gore, J. (2016). Reassembling teacher professional development: The case for Quality Teaching Rounds. *Teachers and Teaching: Theory and Practice*, 23(3), 352-366.
- Boylan, M., Coldwell, M., Maxwell, B., & Jordan, J. (2018). Rethinking models of professional learning as tools: a conceptual analysis to inform research and practice. *Professional Development in Education*, 44(1), 120-139.
- Bredeson, P.V. (2002). The architecture of professional development: Materials, messages and meaning. *International Journal of Educational Research*, 37, 661-675.
- Brown, S., Edmonds, S., & Lee, B. (2001). *Continuing professional development: LEA and school support for teachers*. Berkshire: National Foundation for Educational Research.
- Bryk, A.S. (2015). 2014 AERA Distinguished Lecture: Accelerating how we learn to improve. *Educational Researcher*, 44(9), 467-477.
- Bubb, S., & Earley, P. (2010). *Helping staff develop in schools*. London, UK: Sage.
- Casteel, C.J., & Ballantyne, K.G. (Eds.) (2010). *Professional development in action: Improving teaching for English learners*. Washington, DC: National Clearinghouse for English Language Acquisition.
- Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, 18(8), 947-967.

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

- Coburn, C.E., & Penuel, W.R. (2016). Research-practice partnerships in education: Outcomes, dynamics and open questions. *Educational Researcher*, 45(1), 48-54.
- Coldwell, M., & Simkins, T. (2011). Level models of continuing professional development evaluation: A grounded review and critique. *Professional Development in Education*, 37(1), 143-157.
- Cornman, S.Q., & Zhou, L. (2016). *Revenues and expenditures for public elementary and secondary education: School year 2013-14 (fiscal year 2014)*. Washington, DC: National Center for Education Statistics.
- CUREE [Centre for the Use of Research and Evidence in Education]. (n.d.). *Evaluation of CPD providers in England 2010-2011: Report for CPD providers*. Coventry, UK: Author.
- Darling-Hammond, L., Wei, R.C., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad*. Retrieved from <https://learningforward.org/docs/pdf/nsdcstudy2009.pdf>
- Desimone, L.M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181-199.
- Earley, P., & Porritt, V. (Eds.) (2010). *Effective practices in continuing professional development: Lessons from schools*. London, UK: Institute of Education.
- 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.
- Evans, L. (2011). The 'shape' of teacher professionalism in England: Professional standards, performance management, professional development and the changes proposed in the 2010 White Paper. *British Educational Research Journal*, 37(5), 851-870.

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

- Evans, L. (2014). Leadership for professional development and learning: Enhancing our understanding of how teachers develop. *Cambridge Journal of Education*, 44(2), 179-198.
- Fishman, B.J., Marx, R.W., Best, S., & Tal, R.T. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education*, 19(6), 643-658.
- Flecknoe, M. (2002). Measuring the impact of teacher professional development: Can it be done? *European Journal of Teacher Education*, 25(2-3), 119-134.
- Fraser, C., Kennedy, A., Reid, L., & Mckinney, S. (2007). Teachers' continuing professional development: Contested concepts, understandings and models. *Journal of In-service Education*, 33(2), 153-169.
- Garet, M.S., Porter, A.C., Desimone, L., Birman, B.F., & Yoon, K.S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915-945.
- General Teaching Council for Scotland. (n.d.). *Using evidence of impact and examples of professional learning journeys*. Retrieved from <http://www.gtcs.org.uk/professional-update/professional-learning-journey-using-evidence-of-impact.aspx>
- Goodall, J., Day, C., Lindsay, G., Muijs, D., & Harris, A. (2005). *Evaluating the impact of continuing professional development (CPD)*. Retrieved from <http://webarchive.nationalarchives.gov.uk/20130402123207/https://www.education.gov.uk/publications/eOrderingDownload/RR659.pdf>
- Gore, J.M., & Gitlin, A.D. (2004). [RE]Visioning the academic-teacher divide: Power and knowledge in the educational community. *Teachers and Teaching: Theory and Practice*, 10(1), 35-58.
- Grossman, T., & Hirsch, E. (2009). *State policies to improve teacher professional development*. Washington, DC: NGA Center for Best Practices.

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

- Guskey, T.R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin Press.
- Guskey, T.R. (2002). Does it make a difference? Evaluating professional development. *Educational Leadership*, 59(6), 45-51.
- Guskey, T.R. (2014). Measuring the effectiveness of educators' professional development. In L.E. Martin, S. Kragler, D.J. Quatroche, & K.L. Bauserman (Eds.), *Handbook of professional development in education: Successful models and practices, PreK-12* (pp. 447-466). New York, NY: Guilford.
- Harris, A., Day, C., Goodall, J., Lindsay, G., & Muijs, D. (2006). What difference does it make? Evaluating the impact of continuing professional development in schools. *Scottish Educational Review*, 37, 91-99.
- Haslam, M.B. (2010). *Teacher professional development evaluation guide*. Washington, DC: National Staff Development Council.
- Hill, H.C., Beisiegel, M., & Jacob, R. (2013). Professional development research: Consensus, crossroads, and challenges. *Educational Researcher*, 42(9), 476-487.
- Hochberg, E.D., & Desimone, L.M. (2010). Professional development in the accountability context: Building capacity to achieve standards. *Educational Psychologist*, 45(2), 89-106.
- Hunzicker, J. (2011). Effective professional development for teachers: A checklist. *Professional Development in Education*, 37(2), 177-179.
- Jackson, A., & Burch, J. (2016). School Direct, a policy for initial teacher training in England: Plotting a principled pedagogical path through a changing landscape. *Professional Development in Education*, 42(4), 511-526.
- James, M. (2012). Growing confidence in educational research: Threats and opportunities. *British Educational Research Journal*, 38(2), 181-201.

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

- Kelleher, J. (2003). A model for assessment-driven professional development. *Phi Delta Kappan*, 84(10), 751-756.
- Kennedy, A. (1997). The connection between research and practice. *Educational Researcher*, 26(7), 4-12.
- Kennedy, A. (2005). Models of continuing professional development: A framework for analysis. *Journal of In-Service Education*, 31(2), 235-250.
- Kennedy, A. (2014). Understanding continuing professional development: The need for theory to impact on policy and practice. *Professional Development in Education*, 40(5), 688-697.
- Killion, J. (2003). 8 smooth steps: Solid footwork makes evaluation of staff development programs a song. *Journal of Staff Development*, 24(4), 14-26.
- Killion, J. (2008). *Assessing impact: Evaluating staff development*. Thousand Oaks, CA: Sage.
- King, F. (2014). Evaluating the impact of teacher professional development: An evidence-based framework. *Professional Development in Education*, 40(1), 89-111.
- King, J.A. (2002). Building the evaluation capacity of a school district. *New Directions for Evaluation*, 93, 63-80.
- Lawrenz, F., Thomas, K., & Clarkson, L.C. (2008). Evaluation capacity building in the schools: Administrator-led and teacher-led perspectives. *Canadian Journal of Program Evaluation*, 23(3), 61-82.
- Lauer, P.A., Christopher, D.E., Firpo-Triplett, R., & Buchting, F. (2014). The impact of short-term professional development on participant outcomes: A review of the literature. *Professional Development in Education*, 40(2), 207-227.
- Le Fevre, D.M. (2014). Barriers to implementing pedagogical change: The role of teachers' perceptions of risk. *Teaching and Teacher Education*, 38, 56-64.

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

- Ling, L. M., & Mackenzie, N. (2015). An Australian perspective on teacher professional development in supercomplex times. *Psychology, Society, and Education*, 7(3), 264-278.
- Lingard, R., Sellar, S., & Lewis, S. (2017). Accountabilities in schools and school systems. *Oxford Research Encyclopedia of Education*. Retrieved from <http://education.oxfordre.com/view/10.1093/acrefore/9780190264093.001.0001/acrefore-e-9780190264093-e-74>
- Mitchem, K., Wells, D., & Wells, J. (2003). Using evaluation to ensure quality professional development in rural schools. *Journal of Research in Rural Education*, 18(2), 96-103.
- McChesney, K. R. (2017). *Investigating teachers' experiences of professional development within a major education reform in the emirate of Abu Dhabi*. Thesis (PhD). Curtin University. Retrieved from <https://espace.curtin.edu.au/handle/20.500.11937/57566>
- McChesney, K. R., & Aldridge, J. M. (2018). A new tool for practitioner-led evaluation of teacher professional development. *Teacher Development* (advance online publication). doi:10.1080/13664530.2017.1418418
- McCormick, R. (2010). The state of the nation in CPD: A literature review. *The Curriculum Journal*, 21(4), 395-412.
- Muijs, D., & Lindsay, G. (2008). Where are we at? An empirical study of levels and methods of evaluating continuing professional development. *British Educational Research Journal*, 34(2), 195-211.
- Mullins, D., Lepicki, T., & Glandon, A. (2010). *A professional development evaluation framework for the Ohio ABLE system*. Retrieved from http://uso.edu/network/workforce/able/reference/development/PD_Eval_Framework_Report.pdf
- New Zealand Ministry of Education. (n.d.). *Professional learning and development*. Retrieved from <http://services.education.govt.nz/pld/>

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

Ofsted. (2006). *The logical chain: Continuing professional development in effective schools*.

Retrieved from

<http://dera.ioe.ac.uk/5999/1/The%20logical%20chain%20continuing%20professional%20development%20in%20effective%20schools%20%28PDF%20format%29.pdf>

Ofsted. (2010). *Good professional development in schools: How does leadership contribute?*

Retrieved from

<http://dera.ioe.ac.uk/1109/1/Good%20professional%20development%20in%20schools.pdf>

Opfer, V.D. (2016). *Conditions associated with teacher professional development and its impact on instruction in TALIS 2013*. Paris, France: OECD.

Opfer, V.D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research, 81*(3), 376-407.

OECD [Organisation for Economic Co-operation and Development]. (2014). *TALIS 2013 results: An international perspective on teaching and learning*. Paris: Author.

OECD [Organisation for Economic Co-operation and Development]. (2016). *What makes a school a learning organisation? A guide for policy makers, school leaders and teachers*. Paris: Author.

Pedder, D., & Opfer, V.D. (2010). Planning and organisation of teachers' continuous professional development in schools in England. *The Curriculum Journal, 21*(4), 433-452.

Porritt, V., Spence-Thomas, K., & Taylor, C. (2017). Leading professional learning and development. In P. Earley & T. Greany (Eds.), *School leadership and education system reform* (pp. 121-130). London, UK: Bloomsbury.

Puma, M., & Raphael, J. (2001). *Evaluating standards-based professional development for teachers: A handbook for practitioners*. Washington, DC: Urban Institute Education Policy Center.

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

- Regional Education Laboratory Southeast. (2009). *Making decisions about teacher professional development: Practices in 8 Alabama school districts*. Retrieved from http://www.educationalpolicy.org/publications/pubpdf/EBE_436_Eval%20PD_AL.pdf
- Robinson, M., Walker, M., Kinder, K., & Haines, B. (2008). *Research into the role of CPD leadership in schools*. Retrieved from <https://www.nfer.ac.uk/publications/pdl01>
- Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). New York, NY: Free Press.
- Schoenfeld, A.H. (1999). Looking toward the 21st century: Challenges of educational theory and practice. *Educational Researcher*, 28(7), 4-14.
- Shaha, S.H., Lewis, V.K., O'Donnell, T.J., & Brown, D.H. (2004). Evaluating professional development: An approach to verifying program impact on teachers and students. *The Journal of Research in Professional Learning*, 1(1), 1-18.
- Sleeter, C. (2014). Toward teacher education research that informs policy. *Educational Researcher*, 43(3), 146-153.
- Snow, C.E. (2015). 2014 Wallace Foundation Distinguished Lecture: Rigor and realism: Doing educational science in the real world. *Educational Researcher*, 44(9), 460-466.
- Snow, C.E. (2016). The role of relevance in education research, as viewed by former presidents. *Educational Researcher*, 45(2), 64-68.
- Soine, K.M., & Lumpe, A. (2014). Measuring characteristics of teacher professional development. *Teacher Development*, 18(3), 303-333.
- Speck, M., & Knipe, C. (2005). *Why can't we get it right?: Designing high-quality professional development for standards-based schools* (2nd ed.). Thousand Oaks, CA: Sage.
- Sumsion, J., Brownlee, J.L., Ryan, S., Walsh, K., Farrell, A., Irvine, S., Mulhearn, G., & Berthelsen, D. (2015). Evaluative decision-making for high-quality professional development: Cultivating an evaluative stance. *Professional Development in Education*, 41(2), 419-432.

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

Supovitz, J.A., & Turner, H.M. (2000). The effects of professional development on science teaching practices and classroom culture. *Journal of Research in Science Teaching*, 37(9), 963-980.

TDA [Training and Development Agency for Schools]. (2007). *Supporting the leadership of continuing professional development in schools*. Retrieved from http://webarchive.nationalarchives.gov.uk/20100202130439/http://www.tda.gov.uk/upload/resources/pdf/cpd/supporting_leadership_cpd.pdf

TDA [Training and Development Agency for Schools]. (2010). *Impact evaluation—a model: Guidance and practical examples*. Manchester, UK: Author.

The Teaching Council. (2016). *Cosán: Framework for teachers' learning*. Kildare, Ireland: Author.

Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration*. Wellington; New Zealand Ministry of Education.

TNTP [The New Teacher Project]. (2015). *The mirage: Confronting the hard truth about our quest for teacher development*. Retrieved from https://tntp.org/assets/documents/TNTP-Mirage_2015.pdf

Trygstad, P.J., Banilower, E.R., Smith, P.S., & Nelson, C.L. (2014). *New instruments for studying the impact of science teacher professional development*. Chapel Hill, NC: Horizon Research.

UK Department for Education. (2016). *Standard for teachers' professional development: Implementation guidance for school leaders, teachers, and organisations that offer professional development for teachers*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/537031/160712_-_PD_Expert_Group_Guidance.pdf

TEACHER PROFESSIONAL DEVELOPMENT EVALUATION

Vanderlinde, R., & van Braak, J. (2010). The gap between educational research and practice:

Views of teachers, school leaders, intermediaries and researchers. *British Educational Research Journal*, 36(2), 299-316.

Wayne, A.J., Yoon, K.S., Zhu, P., Cronen, S., & Garet, M.S. (2008). Experimenting with teacher professional development: Motives and methods. *Educational Researcher*, 37(8), 469-479.

Webster-Wright, A. (2009). Reframing professional development through understanding authentic professional learning. *Review of Educational Research*, 79(2), 702-739.

Whitehurst, G.J. (2012). The value of experiments in education. *Education Finance and Policy*, 7(2), 107-123.

Wilson, S.M., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. *Review of Research in Education*, 24(1), 173-209.

Yoon, K.S., Duncan, T., Lee, S.W., Scarloss, B., & Shapley, K.L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement*.

Retrieved from https://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_2007033.pdf