

Evaluating the impact of teacher professional development: an evidence-based framework

Journal:	<i>Professional Development in Education</i>
Manuscript ID:	Draft
Manuscript Type:	Original Article
Keywords:	Impact, Evaluation, Framework, Professional Development
Abstract:	<p>Does teacher professional development make a difference? How do we know? While researchers and policy makers acknowledge that teacher professional development (PD) needs to be assessed and evaluated there is often little clarity as to how this can be achieved. Evaluation of teacher PD has been described as the weak link in the PD chain despite it being linked with improved PD experiences and pupil outcomes. This may be due to a lack of skills and tools to carry out such evaluations or indeed it may be linked to how PD is conceptualised. This article explores extant literature and models of evaluation, revealing gaps in existing evaluation frameworks resulting in the development of a provisional PD evaluation framework for use in a study which set out to formally evaluate the impact of a PD initiative on teachers' professional learning in five urban primary disadvantaged schools in the Republic of Ireland. Following application within this study the framework was critiqued and revised resulting in a new PD Impact Evaluation Framework which has subsequently been further tested to add to its robustness and may support teachers, researchers and policy makers to carry out systematic and focused evaluations of teacher PD.</p>

SCHOLARONE™
Manuscripts

Evaluating the impact of teacher professional development: an evidence-based framework

Introduction

The demands of a globalised economy, and the perceived need to perform well in international measures of educational effectiveness (such as PISA league tables), ensure that despite global recession many governments continue to invest in teachers’ professional development as a means of enhancing pupil outcomes. This is certainly the case in the Republic of Ireland, where this research was conducted, and where major professional development initiatives are being promoted despite major cuts in public spending generally and education in particular. However, researchers and policy makers continue to struggle to establish whether professional development (PD) makes a difference (Rhodes *et al.*, 2004; CUREE, 2008). Traditionally measuring impact of teacher PD has focused largely on teacher satisfaction and ignored impact on teacher learning, use of new practices, pupils’ outcomes and/or value for money (Rhodes *et al.*, 2004; O’Sullivan, 2011). Given this ongoing investment, in current straitened times, and an international trend towards accountability, there is an increasing emphasis and need to evaluate the impact of teacher PD.

As indicated, this is reflected in the Irish context where in July 2011, the Department of Education and Skills (DES), launched the ‘National Strategy to Improve Literacy and Numeracy among Children and Young People 2011 -2020’, hereafter referred to as ‘the Strategy’. Within the Strategy there is a significant commitment to teacher PD aimed at improving educational standards. The need to assess and evaluate the impact of this PD is also reflected within the Strategy.

While evaluating the impact of teacher PD forms part of education policy in many countries, it tends to be problematic (Rhodes *et al.*, 2004; CUREE, 2008). Earley and Porritt (2010) argue that the difficulty may lie with a lack of appropriate expertise and inadequate tools to carry out such evaluations. Therefore the purpose of this paper is to begin to address this need through presenting an evidence-based framework, which sought to formally evaluate the impact of a PD initiative on teachers’ professional learning. Within this study teachers’ professional learning is defined as the growth of teacher expertise leading to a change in practice which results in improved pupil learning (New South Wales (NSW) Institute of Teachers, 2007).

This paper presents a framework for evaluating the impact of professional development on teachers' professional learning. It does so by reviewing the extant literature, from which the *Provisional Framework* was developed. The paper outlines the context of the study, and the methodology used to apply the framework. In light of this application the paper sets out how the framework was revised and how it can be both developed and applied in a professional context with teachers.

Context

The study was carried out in five urban disadvantaged primary schools in the ROI as categorised by the Social Inclusion Section of the DES using the DEIS (Delivering Equality of Opportunity in Schools) Banding categorization. These same schools were the focus of a study, three years previously, which focused on measuring the impact of a literacy initiative on pupil outcomes. The schools involved were selected from 19 schools that responded to an advertisement in the Irish National Teachers' Organisation (INTO) (teacher union) magazine inviting schools to engage in a literacy initiative in which they would be funded and supported by the INTO. Funding provided all materials, the input of project facilitators and release time from school for teachers to engage with the PD initiative. Additional support was provided in terms of two school visits from a project facilitator during the implementation period and access to support via email and telephone. The initiative centred on pupils in 3rd class (average age 9) and involved a special educational needs (SEN) teacher and a classroom teacher working collaboratively within the mainstream classroom to facilitate Peer Tutoring (PT) (Topping, 1988; Butler, 1999) for literacy for thirty minutes a day, four days a week, over an eight to ten-week period. PT involves pupils reading in mixed ability pairs in the role of tutor and tutee in a bid to improve their reading accuracy and fluency. The initiative appeared to impact pupil achievement, with findings at the time suggesting an overall average improvement of 12.7 months in reading accuracy for pupils (n=116), high levels of pupils' enjoyment and teachers' willingness to sustain the practices beyond the initial input (Author and Co-author, 2009). While these findings are not part of this study, the schools involved are, as this study involved a return to the same five schools to explore the impact of the original PD initiative on teachers' professional learning three years on. The rationale for this study came from the literature which has not only identified a paucity of research measuring the impact of PD generally but also little evidence of sustainability of teaching practices in the longer-term despite sustainability of practices being pivotal for school improvement

(Baker *et al.*, 2004; Priestley *et al.*, 2011). This paper is a report of how the longer-term impact of a professional development initiative was formally evaluated using an evidence-based framework.

The aims of this paper are fourfold: first, to explore the challenges associated with evaluating PD; second, to analyse the role and limitations of previous evaluation models; third to demonstrate how a *Provisional Framework* was devised for use in this study; and fourth to set out how this framework was adapted following implementation to result in the evidence-based *PD Impact Evaluation Framework* presented here. Therefore the focus of this paper is on the design and development of an evaluation framework which was used to assess the impact of a PD initiative within the context of the above study.

Evaluating the impact of PD

Evaluating the impact of PD has been reported as the weakest link in the PD chain (Ofsted, 2006) despite it being described as ‘a learning tool that improves the quality of both the CPD activity and the outcomes achieved’ (Earley and Porritt, 2010: 147). The term CPD (*continuing professional development*) is used in some countries to refer to PD which takes place after initial teacher education while the term PD can be often associated with initial teacher education. Given that the focus of this paper is on measuring impact of teacher PD which applies across the continuum of teacher education, the term PD will be used to reflect this continuum.

Evaluating PD is often neglected (Earley and Bubb, 2004), problematic (Rhodes *et al.*, 2004; CUREE, 2008) and seldom carried out in a ‘systematic and focused manner’ (Muijs and Lindsay, 2008: 196). Generally, it appears to be ‘instinctive, pragmatic and without explicit reference to clearly defined learning outcomes for teachers or students’ (Opfer *et al.*, 2010: 10). Arguably, this may be due to either the contested notion of what constitutes teacher PD, or not having the ‘experience, skills and tools’ to do so (Earley and Porritt, 2010: 6). If PD is viewed in the traditional sense of ‘inputs’, such as courses, rather than the ‘actual development of knowledge and expertise (outcomes)’ then this may also impact on its evaluation (Bubb and Earley, 2008: 5). This concept of PD reflects the definition by the NSW Institute of Teachers (2007: 3) who describe it as ‘the processes, activities and experiences that provide opportunities to extend teacher professional learning’ which ultimately focuses

on pupil outcomes. Interestingly the DES (2011: 37) in the ROI has stipulated within the Strategy that 'PD courses be adequately assessed and evaluated' which raises again the question of the relationship between PD and 'courses'. With the introduction of the Strategy, a commitment to undertake PD is now mandatory for teachers in the ROI; this is new for Irish teachers for whom engagement with PD was previously voluntary. It is of little surprise then that a policy for assessing and evaluating the impact of this PD, which the government are significantly investing in, is a commitment in the Strategy. What is much less clear is how to do this.

PD Evaluation: competing models

To formally evaluate the impact of the PD experience in this study, an exploration of the literature was undertaken to identify a suitable framework to facilitate this. This included an exploration of a wide range of literature (much of which has been published in this journal) and which have been associated with trying to establish how to measure impact. A number of evaluation 'models' were identified and these are presented and critiqued here. The first example is provided by Kirkpatrick (1959) who devised an evaluation framework for evaluating business training programmes, consisting of four levels:

1. Participant reaction
2. Participant learning
3. Participant behaviour
4. Organisation results

A framework more specific to education was that by Stake (1967) who identified: 'antecedents' (how things were before the programme); 'transactions' (what occurred during the programme) and 'outcomes' (results of the programme). The challenge with this framework lies in the difficulty establishing the cause and effect relationship between a programme and its outcomes; as there may be many influencing variables from within or outside of the programme, for example, teacher personal development or the programme itself (Holton, 1996; Frost and Durrant, 2003; Stevenson, 2008).

Thomas Guskey (2002) developed Kirkpatrick's framework and introduced five levels of professional development evaluation for use in a school context; the additional level titled *Organisation support and change*.

1. Participants’ reactions
2. Participants’ learning
3. Organisation support and change
4. Participants’ use of new knowledge and skills
5. Students’ learning outcomes.

(Guskey, 2002: 47)

Students’ learning outcomes replaced Kirkpatrick’s *Organisational results* while participant’s use of new knowledge and skills replaced Kirkpatrick’s *Participant’s behaviour*. Guskey’s (2002) model was further developed by Bubb and Earley (2010) and comparisons and additional levels can be seen in Table 1.

Table 1: Comparison of evaluation models about here

The first three levels in Bubb and Earley’s (2010) model seek to facilitate PD planning, as evidence suggests that planning PD in advance can improve the outcomes for teachers and students (Earley and Porritt, 2010). These planning levels have been adapted from Guskey’s (2002) suggestion to reverse his five levels of evaluation for PD planning. They also reflect Stake’s (1967) notion of *antecedents* in that they set out to establish how things were before the commencement of PD. Establishing a baseline in this way enables participants to plan the way forward and chart their development and progress over time, unlike Kirkpatrick’s (1959) model which seems to focus solely on evaluation of training.

Guskey’s (2002) level 1 is concerned with *Participants’ Reactions* and initial satisfaction with various aspects of the PD experience which is reflective of the level of current evaluations used by many engaging in PD evaluations. Meanwhile, level 2 centres on *Participants’ Learning* to ascertain if the intended knowledge and skills were realised by participants. Bubb and Earley (2010) also have similar levels (4 and 5) in their model, namely *The Experience* and *Learning* respectively. These focus on the same issues with the significant added dimension of a focus on ‘attitudes’ - acquired or enhanced. This is not explicit in Guskey’s (2002) model as he appears to assert that it is the practice or experience that will shape teachers’ attitudes and beliefs (Guskey, 2005) even though others argue that this is a cyclical process with teachers entering the cycle at various starting points (Opfer *et al.*, 2010). This raises an issue surrounding the concept of levels within evaluation or impact models especially with Guskey’s assertion (2000) that these levels are successive and interrelated with one level building on the other and indeed success at each level being

necessary for subsequent levels. However Coldwell and Simkins, (2011) point out that level 3 of Guskey's (2002) model, *Organisation Support and Change*, which identifies contributing factors to support implementation and change are not a consequence of level 2, *Participant's Learning*. The same could be said of Bubb and Earley's (2010) use of levels, which also suggests a hierarchical model. This hierarchy between levels has been critiqued by some (Holton, 1996; Coldwell and Simkins, 2011) who dispute the successive notion of levels and argue for a more intricate and nuanced model of evaluation that will encompass the complex nature of all the variables that affect teacher engagement with PD and teacher and pupil outcomes. Such an approach seeks to overcome the problems associated with an over-simplified sequentialism in the models that depend on using levels.

Both models Guskey's and Bubb and Earley's models agree on the need for *organisation support* for implementation and change; this may be in the format of a change agent, leadership, policies or resources. Bubb and Earley (2010: 61) have *Organisational Support* level 6 which identifies 'how the school helps (or hinders) the person using their new learning in their job'. Guskey's model (2002) also explores impact on the organisational climate and procedures at this level. However there is a need to explore the more complex relationships that exist and influence teacher engagement with PD. For example one missing component from both models is that of collaboration amongst teachers centred on developing a shared vision of pedagogy and pedagogic content knowledge (Darling-Hammond, 1997; Smith, 2007), which is seen as the cornerstone for implementing and sustaining practices for school improvement.

Level 4 of Guskey's (2002) model concentrates on evaluating *Participants' use of new knowledge and skills* and aims to measure the degree and quality of implementation of new practices. Bubb and Earley's (2010: 61) model takes a similar but slightly different slant in that it focuses on Into practice - the degree and quality of change in level 7, to account for new products (such as policies), new processes (new or improved systems) and staff outcomes. They drew upon the work of Frost and Durrant (2003) and the TDA (2006) to explore different levels of impact on staff: classroom practice, personal capacity and interpersonal capacity. However, there is little clarity as to how to measure the degree of implementation of new practices with either model. Given that the literature is replete with calls for teachers' levels of engagement with practices to be at a deep, conceptual level for sustainability of practices (see for example Baker *et al.*, 2004; Bolam *et al.*, 2005; Priestley *et al.*, 2011), it is essential to have a clear framework for measuring this. Hall and Hord (1987)

provides such a framework for evaluation of implementation which outlines eight ‘Levels of Use’ (LoU) of implementation (see Table 2).

*Table 2 Levels of Use of the Innovation **about here***

For sustainability of practices teachers need to have the conceptual knowledge or deep learning related to the practice as evidenced at the routine and refined levels of use. Interestingly, in Baker *et al.*, (2004) it is argued that a considerable number of teachers never progress beyond the routine level of understanding and use. However, teachers may need support to progress through these levels (Hall and Hord, 1987). This support may be in the form of creating collaborative cultures of teaching and learning (Hodkinson and Hodkinson, 2005) such as professional learning communities (PLCs) or other forms of collaboration; team teaching practices; peer coaching; or collaborative consultation, to enhance the system’s overall capacity (O’Sullivan, 2011). Such collaborative cultures may in turn impact on other beneficiaries (Stevenson, 2008) where the impact of PD is extended to other teachers and pupils in a process described by Earley and Bubb as ‘cascading’ (2004: 84). This features in Bubb and Earley’s (2010) model at levels 9-12 and is essential if practices within schools are to survive despite staff turnover. To facilitate such cascading of practices, school re-culturing may be required recognising that ‘the real agenda [for school improvement] is changing school culture not single innovations (Stoll and Fink, 1996: 45-46).

Notwithstanding Bubb and Earley’s (2010) acknowledgment of teacher attitudes as important in the implementation and change process, the literature emphasizes the importance of winning teachers’ ‘hearts and minds’ (Evans, 2010). Some concerns have been expressed regarding little evidence of changes in teachers’ beliefs and values (Gleeson and O’Donnabháin, 2009; Opfer *et al.*, 2010) and therefore it is important to explore this as part of Bubb and Earley’s (2010) level 7 Into practice-‘Degree and quality of change’.

Finally, Level 5 of Guskey’s (2002) model and level 8 of Bubb and Earley’s (2010) model focuses on *Student learning outcomes*. A particular strength of both models is the inclusion of impact on students’ cognitive, affective and psychomotor skills, which echoes the work of Bloom (1956).

An analysis of evaluation models and the relevant literature discussed above revealed a number of gaps in existing frameworks (for example, Kirkpatrick, 1959; Stake, 1967; Guskey, 2002 and Bubb and Earley, 2010). As a consequence the framework for this study

sought to deliberately build on previous work, in particular Guskey (2002) and Bubb and Earley's (2010) ground-breaking work, but to develop it further. This led to a synthesis of findings and, for the purposes of this study, the development of a new *Provisional Framework* (Figure 1). This new provisional model sought to acknowledge the strengths of previous models whilst also addressing limitations and gaps especially in the area of collaborative practices; which are heavily endorsed in the literature, as essential components for teacher learning, sustainability of practices and whole school change and development. This *Revised Framework* also included teachers' deep learning, teacher commitment and teacher ownership of practices which were also highlighted in the literature. As this study focused on the people who did engage (users), with the initiative at different LoU (Hall and Hord, 1987) the 'Nonusers' levels were not included on the framework. Furthermore, an amalgamation of the refined and integrated LoU similar to that used by Baker *et al.*, (2004) was employed for this study, as the initiative being evaluated was collaborative by nature and therefore separate headings were considered inappropriate.

Additionally, more emphasis was placed on teacher attitudes and beliefs given their significance as being central to the change process (Opfer *et al.*, 2010). This *Provisional Framework* guided data collection and analysis as outlined in the methodology. It was critiqued following application in this study and refined resulting in an evidence-based *PD Impact Evaluation Framework* which may help address the need for a tool or common conceptual framework (Desimone, 2009) for evaluating the impact of teacher PD at a time when demonstrating the impact of PD, and indeed its value for money, are becoming increasingly important.

Insert Figure 1: Provisional Professional Development Impact Evaluation Framework about here

Methodology

A multiple case study approach was used to carry out this qualitative research initiative, which sought to explore the impact of a PD initiative, on teachers' learning in five urban disadvantaged schools in the ROI, using the devised *Provisional Framework* (Figure 1). Within each of the five schools the sampling used was purposive given that it was a follow-up from an initial research project in which there was a class teacher, SEN teacher and principal from five schools. However due to staff turnover and the flexible nature of case study research, it was possible to interview people in those schools who had subsequently

become involved in the initiative, to facilitate an exploration of the ‘legacy’ (Author, 2012) of the PD initiative in these schools.

A conceptual framework, which is the ‘current version of the researcher’s map of the territory being investigated’ (Miles and Huberman, 1994: 20), was used for directing the data collection. The framework used was initially based on Miles and Huberman’s (1994: 18) ‘Conceptual Framework for a Study of the Dissemination of Educational Innovations’ and was adapted for use in this study. The extensive literature review and analysis of existing evaluation models led to the development of a *Provisional Framework* which incorporated all aspects from the original map and reflected the improved understanding of the ‘terrain’ (Miles and Huberman, 1994: 18). This subsequently guided the research design and process. The contents are largely based on the work of Guskey (2002) and Bubb and Earley (2010) and supplemented from various other sources, for example, Hall and Hord (1987). The sections in grey reflect the ‘study propositions’ from the literature which helped direct the research design, even though the propositions that were outlined at the beginning of the research may no longer have been valid following data analysis (Yin, 2009: 34). Therefore, the propositions included in this *Provisional Framework*, which were predicated on the literature, were critiqued as part of this research.

Given the context of this study the *Provisional Framework* deviated from Bubb and Earley’s (2010) levels 1-3 which emphasised the importance of establishing baseline data and planning the PD experience. As this study is based on previous research baseline data for teachers’ learning did not exist. At the time of the initial PD it was never intended to return to this cohort of teachers. While this is a limitation, it was possible to use retrospective recollections from teachers to explore the impact of the PD initiative three years on. The initial concerns about teachers’ ability to remember how things evolved over time were allayed through probing for details during interviews where most teachers looked back and saw their learning journey clearly.

The *Provisional Framework* (Figure 1) informed the research questions and subsequent interview questions with the 20 participants from the five schools. It was therefore devised, applied, tested and critiqued in the current study. While this is a small-scale case study the emphasis was on a trustworthy study (Hammersley, 2007) and ‘replication’ can be claimed when two or more cases are shown to support the same theory (Yin, 1994: 31). Results from each of the five case studies in this research provide extensive evidence to back up findings,

and conclusions were reached only after the findings were tested or confirmed through checking rival explanations, variables, and feedback from interviewees where necessary (Miles and Huberman, 1994; Punch, 2009; Yin, 2009).

Applying and testing the framework

The framework was operationalised for use in this study through linking a set of research questions with each section of the framework. The questions sought to: focus on short-term and long-term impact in an effort to fill the research gap relating to sustainability of new practices in schools; explore the factors that helped or hindered the development and sustainability of PD practices and; assess impact using the evaluation framework devised from the extant literature.

The sections of the framework dealing with *The Experience, Learning and Organisation Support* focused on short-term impact of the PD initiative. The section on *Into Practice* is the largest section and focused on the long-term impact, measured in terms of process, product and staff outcomes. Teachers' perceptions of *pupils' outcomes* were evaluated at various levels as highlighted on the framework. Whether or not there was a cascading of knowledge and practices is explored in the *Cascading* section on the framework. Following application of the framework, (through the research and interview questions) and data analysis, it was further refined and those refinements are reported and discussed later in this paper.

A rigorous, transparent and systematic approach to data analysis was employed (Robson, 2002; Bryman, 2004). The use of a software package, QSR NVivo 8, facilitated the collection and storage of data in an organised manner. Interviews were transcribed as soon as possible after being carried out so that contextual cues and nonverbal cues were not lost (Silverman, 2000). Descriptive or topic codes were used on initial readings. Inductive data analysis was used where the codes were not predetermined (Bryman, 2004; Gray, 2004). Interestingly some of the initial codes were quite reflective of the headings on the *Provisional Framework*. This alignment between emerging codes and the PD framework headings helped in the critique of the framework itself for use as a tool for evaluating PD. On subsequent readings the relationship between some codes appeared and memos developed (Punch, 2009). This led to a second round of codes and so the iterative process of data analysis continued, with explanations devised for each code along the way. This also provided a clear audit trail to assess procedures and findings (Lincoln and Guba, 1985). These explanations were then used

to code the data on subsequent readings and led to data reduction through merging and omission of certain codes with similar meanings, for example *pedagogy* and *knowledge-conceptual level* were omitted and merged under *use of knowledge and skills –routine level* and *refined/Integrated* level as it is reflective of having a conceptual knowledge of the practice. Analysis continued in this way with data reduction, data display through the revision of codes and definitions, and drawing and verifying conclusions (Miles and Huberman, 1994). The final round of codes resulted in the *Provisional Framework* (Figure 1) being developed to reflect the codes arising from the data analysis as outlined in the next section.

Findings

While the study focused on exploring the short-term and long-term impact of a PD initiative and the factors that helped or hindered the development and sustainability of the initiative, the focus of this paper is on the design and development of an evaluation framework which was used to assess the impact of a PD initiative within the context of the study. Therefore the findings as reported here relate to the design and development process of the framework.

Following application and testing of the *Provisional Framework* as outlined above, its suitability for evaluation of the PD initiative was assessed through data analysis as described. Overall, findings indicated that the framework was ‘fit for purpose’ in evaluating the impact of the PD initiative in this study.

While most of the headings on the framework worked well, some were merged and others renamed, leading to further development and refinement of the framework (Figure 2). These developments will be now explained in detail under each section of the *Revised Framework*.

Insert: Figure 2: The *Revised PD Impact Evaluation Framework*

The Experience

This section of the framework was concerned with teachers’ satisfaction with the initial PD off-site training experience.

Learning

This section was intended to focus on new or enhanced skills, attitudes and knowledge in the short-term, that is, the initial eight to ten-week implementation of the literacy initiative. However given the nature of this PD initiative, which comprised of one full day's PD followed by two schools visits and email / telephone support over the implementation period, it became clear that this section of the framework could not be used to assess learning over this length of time. Data related to this initial implementation period highlighted factors from all other levels of the framework: *Organisation support; Into Practice- Degree of quality and change; Pupil Outcomes* and; *Cascading*. However some participants did refer to the learning on the initial training day and so this *Learning* section remains on the framework but is only suitable to assess new or enhanced knowledge, skills and /or attitudes acquired on the day of a PD experience, activity or course. These two sections are quite reflective of much PD evaluation, which centres largely on teacher satisfaction with the PD experience. To explore impact on teachers' learning and pupils' outcomes, responses from participants about the short-term (initial implementation period) and long-term implementation (three years later) of the practice were recorded under the following sections of the framework: *Organisation Support; Into Practice- Degree of quality and change ; Pupil Outcomes; and Cascading*.

Organisation support

Findings from the data showed that *Organisation support* alone was not sufficient to support teachers in their implementation or sustainability of the literacy initiative. The research clearly identified three consistent features of teachers' professional learning for teacher engagement with implementation and sustainability of the practice:

Support: in the form of Leadership, Professional Learning Communities and an Advocate or change-agent. Leadership support was the mechanism through which other supports, such as the development of PLCs and the modelling of practices by an advocate (someone who is driving and supporting the practice), were enabled to develop (Author, 2011).

Initiative Design and Impact: The structure and design of the literacy initiative, along with its success for students were significant factors in its 'legacy' (Author, 2012).

Teacher Agency: Teachers’ openness and willingness and motivation to engage with and sustain the practice were significant in its survival. Underpinning this was teacher agency, that is, teachers acting in intentional ways to ‘shape their own responses to problematic situations’ (Fallon and Barnett, 2009: 12).

Teacher professional learning is a complex process involving the interconnectedness and interdependency of teacher agency, the initiative itself and the pivotal role of support, as emerged from findings in this study. Therefore the section, *Organisation support* was replaced with *Systemic Factors* to account for these significant additional features that may better represent the complex relationships between teachers, PD and outcomes. This is reflective of arguments posited by some who argue that PD is contextual and subjective (Prosser and Trigwell, 1999; Bottery, 2006) and it reflects an underlying ontological position that the reality of the social world is complex and constructed by the participants who engage within it. Understanding the subjective world of human experience may help enhance the impact of PD.

Into Practice – Degree and quality of change

This section was used to explore the impact of the PD initiative on the teachers’ professional learning. This section was largely based on that of Bubb and Earley’s (2010) framework with additions from the literature as indicated by the shaded areas in Figure 1. The data showed evidence of new *products* and *processes* following engagement with the PD experience and it confirmed what Earley and Porritt (2010) stated about *processes* being necessary to enact certain *products*, for example while schools may have had a new policy regarding implementation of the literacy initiative, which came under the heading *products*, certain *processes*, such as timetabling this at the beginning of the year, needed to be put in place to act upon these products. Many other *processes* reported by participants were reflective of collective practices and so were recorded under that heading.

Staff outcome resulted in some additions and mergers of headings. For example, at a *personal* level there were initially two sub-headings as seen in Figure 1: *feelings and thinking related to classroom teaching*, and *beliefs and attitudes towards pupils’ learning*. On second round coding, another level was added: *feelings and thinking related to collaborative practices*. However, following data reduction two of these sub-headings were merged into one: *beliefs*

and attitudes related to classroom teaching and pupils' learning, while the heading of feelings and thinking related to collaborative practices was transferred to the collective heading. The data revealed that teachers' beliefs, values and attitudes were instrumental in the sustainability of the practice and that changes are iterative and can begin with a change in teacher beliefs leading to a change in practice or can begin with a change in practice leading to a change in beliefs. For example, teachers' beliefs about the value of pupils working in pairs, led to further practices involving pairing of pupils. The positive impact on pupils led to sustainability of the practice and also encouraged others to engage with it, as they valued other teachers' opinions about its success and value (Boardman *et al.*, 2005).

Teachers also expressed a sense of confidence and efficacy in relation to the teaching of reading in this way and therefore *teacher efficacy* was transferred to the personal level as it is connected with teachers' beliefs in their power to effect change, with correlations between affect and efficacy (Kitching *et al.*, 2009). Similarly, some teachers' experience of this literacy practice led to changes in beliefs and values about collaborative practices among teachers, which in turn led to the adoption of other collaborative practices. This is reflective of Opfer and colleagues' (2010) cyclical nature of teacher change, which focuses on the interplay between these variables, in favour of Guskey's (2005) model, which argues that change is linear with changes in beliefs following a change in practice. This section, *Personal*, therefore looks at impact in terms of teachers' beliefs, as they may influence teacher efficacy, practices and pupils' outcomes.

In the *Professional* section there was a merger of the existing headings of teachers' knowledge of innovation and use of new and improved knowledge and skills, as these can be described at three levels, *mechanical*, *routine* or *refined/integrated* (Hall and Hord, 1987; Baker *et al.*, 2004), and are reflective of knowledge at procedural and conceptual levels. This heading became the *quality of use and understanding of new and improved knowledge and skills*. However, it was decided to rename the *mechanical* level of understanding and use, to *technical*, as the former suggests something that is automatic or routine whereas *technical* is more concerned with the logistics and organisational or procedural issues as suggested by Hall and Hord (1987) in their descriptors of behaviours at that level.

Findings also indicated a link between data at Hall and Hord's (1987) *routine* levels of quality of use and understanding and that of teachers' knowledge of *pedagogy*. However,

characteristics at the *routine* level show teachers’ conceptual understanding related to the initiative itself, whereas *pedagogy* encompasses more than this; it focuses on enabling pupils’ learning (MacNeill *et al.*, 2005). Furthermore, Pedagogic Content Knowledge (PCK)– that is, ‘knowledge of ways of representing specific subject matter for pupils and an understanding of the difficulties they may face because of their existing conceptions’ (Smith, 2007: 378) – is also an important part of pedagogy, which is not in evidence at Hall and Hord’s (1987) *routine* level. Therefore *routine* was changed to *accepted* levels of understanding and use, as the primary focus seems to be on teachers accepting that the initiative is working well for their pupils and their conceptual knowledge relates to the initiative itself only.

Meanwhile, the *refined/integrated* level of understanding and use which focuses on making changes to enhance the impact on pupils and working with other teachers to provide a collective impact on pupils is somewhat reflective of the data related to evidence of shared pedagogy and PCK, where teachers have collectively generated new knowledge and practices from their experience. However there is an explicit emphasis on PCK and teachers critically engaging with the initiative and therefore this heading was changed to *critical*.

Findings also revealed that some teachers discontinue the changes in practice, which may be as significant as those who sustain changes in some instances. Therefore a *Discontinued* level was inserted to reflect this. While Hall and Hord’s (1987) three levels of non-users were not applicable to this study, they may be of relevance in other situations for assessing impact, and are therefore included in the *Revised Framework* (Figure 2), hereafter referred to as the *PD Impact Evaluation Framework*.

Underpinning all levels of teacher engagement is teacher agency, which may be considered a requirement more than an impact, and therefore the heading of *human agency* was removed from the framework. *Commitment and ownership* was omitted as it forms part of teacher agency, as teachers are showing commitment and ownership when they are acting in intentional ways to enable change.

Moving to the *collective* section of the framework, some headings were collapsed into each other and replaced with the term *cultural* to reflect the way things are done in a school (Norris, 2004), for example the *forms of collaboration* that ensued from this initiative and the *development of PLCs*. *Staff morale* was omitted, as the data was categorised under affective

levels also and therefore a separate heading was not required. *Interpersonal capacity* was merged with PLCs, as data from the interviews was coded under both headings. Many of the concepts associated with Frost and Durrant's (2003) *interpersonal capacity* are similar to those of PLCs, with the latter including the important additional aspects of teachers having collective responsibility for pupils' learning, shared values and vision, and reflective professional enquiry; essential components for enhancing pupils' outcomes and building capacity for school improvement (Hodkinson and Hodkinson, 2005; O'Sullivan, 2011). Therefore they need to be explicitly included and evaluated as part of any evaluation of PD, and not just expected as part of their performance management.

Taking cognisance of *staff outcomes* at *personal*, *professional* and *cultural* levels may help to provide a more comprehensive approach to looking at the depth and degree of teacher understanding and use of new practices, rather than simply acknowledging changes in teachers' practice, skills and knowledge, which previous frameworks seem to do.

Pupils' Outcomes

This section had no changes following data analysis as the findings reflected samples of impact on pupils at each of the levels on Figure 1.

Cascading

This section emerged as highly relevant despite no mention of cascading to adults or pupils in other schools. While this aspect of cascading was not directly relevant to this research, it is important to remember that relying exclusively on site-based learning may lead to lost opportunities for sharing of ideas and resources, less collaboration among teachers from various contexts, less efficient use of outside expertise and less exposure to a broad vision for improvement (Guskey, 1996). Therefore, these remained on the framework, as many PD experiences will occur off-site. However, it was decided to change the term *Cascading* to *Diffusion*, as the former suggests a deliberate, planned, downward movement whereas *Diffusion* is more reflective of the 'rippling' of practices that appeared to occur in this study where teachers willingly engaged with and sustained the practice with the pivotal support of leadership (Author, 2011).

While this study involved using teachers' retrospective recollections, future studies would aim to establish a baseline from which to plan PD experiences and measure impact as

endorsed by Bubb and Earley (2010). Therefore a further three levels were added to reflect Bubb and Earley’s (2008: 61) three levels of ‘baseline picture, goal and plan’, where schools review where they are currently by gathering evidence and then prioritising what they want to achieve, setting a target and then devising plans to get there: in essence, self-evaluating. Following all additions, mergers and omissions of headings the *PD Impact Evaluation Framework* is that in Figure 2. It is hoped that this framework will contribute to the debate on measuring impact of teacher PD through answering the need for a common conceptual framework for evaluation of teacher PD (Desimone, 2009) at a time when policy-makers and researchers are struggling to show that teacher PD makes a difference.

Conclusion

This paper explored how the impact of a PD initiative was evaluated in the context of teachers’ professional learning in five urban disadvantaged primary schools in the ROI. The need for PD to be assessed and evaluated has been highlighted by many (for example Ofsted, 2006; Earley and Porritt, 2010; DES, 2011). However, what is not clear is how to do this. This study provides an evidence-based *PD Impact Evaluation Framework* which may help in providing educationalists with a tool to assess and evaluate teacher PD.

Although previous frameworks for evaluation exist, several gaps were noted in their suitability for exploring the impact of PD. These gaps came under the headings of: *Systemic Factors, Diffusion* and *Staff Outcome* including *Personal, Professional* and *Cultural* impact. Given that the links between teacher PD, pupils’ outcomes and school improvement are not automatic, it is necessary to focus on the processes that facilitate such links, something this new framework takes into consideration through its *Systemic Factors*. These factors acknowledge that change is contextual and influenced by a myriad of factors which can result in change impacting in different ways. Teaching and learning are contextual, and ensuring that PD processes take cognisance of individual professional identities, dispositions, roles and the setting in which teachers work is important to ensure improved outcomes for pupils (Prosser and Trigwell 1999; Hodkinson and Hodkinson, 2005; Bottery, 2006). In this way this framework represents a more complex and nuanced approach to understanding PD and its impact and may be more reflective of the complexity of the social world (Coldwell and

Simkins, 2011). It is important that these are taken into consideration when planning teacher PD experiences with the outcomes in mind.

Another gap in previous frameworks was in the significant role that teacher attitudes and beliefs play in the sustainability of practices, as evidenced in the literature and this study. Therefore these were made more explicit in this new framework under the heading of *Staff Outcome - Personal*. In the *Professional* context the importance of exploring the depth and quality of use and understanding of new and improved knowledge and skills (Hall and Hord, 1987) is acknowledged in the literature and in the findings of this study where teachers' deep levels of learning or conceptual understanding of practices are deemed essential for sustainability of practices (Baker *et al.*, 2005). Another considerable gap in previous frameworks was noted in the area of collaborative practices which are seen as the cornerstone for change. Collaborative professional dialogue and practice are required for school improvement and therefore emphasis on this is included in the *Cultural* section of the new framework. The final change in this new framework is in the area of *Diffusion* which is something to be considered at the outset of PD experiences where schools may consider how new practices may be disseminated to others to enable sustainability of practices for school improvement, whilst recognising that 'diffusion' is both a formal and informal process. Hence, diffusion is considered more apt than cascading in a top down fashion, as it may answer the need to support teachers' openness and willingness for change. Despite this there is an acknowledgement that sometimes a mixture of pressure and support may be needed for change to occur (Stoll and Fink, 1996).

This devised framework may be useful in helping teachers and schools fulfil the need for PD to be adequately evaluated and to move from looking at teacher satisfaction to exploring impact on teacher practices, which in turn aim to enhance pupils' outcomes and school improvement. Significantly for teachers, it can enable them to assess the impact of their own PD, and to know if what they are doing is 'making a difference'. This demonstration of professional responsibility may be a richer form of accountability and one that is more 'fit for purpose' (Collins and Dolan, 2011: 87) than the current emphasis in many countries on over-using standardised tests results as the principal means of judging school improvement. Where schools use this *PD Impact Evaluation Framework* to plan and assess their PD it may help promote individual and collective responsibility for pupils' success. This shared responsibility may in turn help teachers to link PD with school improvement (Pedder *et al.*,

2008) and to promote strong mutual trust within the education system. A more accessible and practitioner-friendly version of the framework is currently being developed so that the framework can be easily used by teachers and school leaders to both plan and evaluate PD activities in schools. It is hoped this version of the framework will help address Bubb and Earley's (2008: 6) call 'to design and test a series of questions for school staff about the quality of learning resulting from the opportunities made available to them'.

References

Baker, S., Gerston, R., Dimino, J. A. and Griffiths, R. (2004) The Sustained Use of Research-Based Instructional Practice: A Case Study of Peer-Assisted Learning Strategies in Mathematics. *Remedial and Special Education*, 25(1), pp. 5–24.

Boardman, A.G., Arguelles, M.E., Vaughn, S., Hughes, M.T. and Klinger, J. (2005) Special Education Teachers' Views of Research-Based Practices. *Journal of Special Education*, 39(3), pp. 168–180.

Bolam, R., McMahon, A., Stoll, L., Thomas, S., Wallace, M., Greenwood, A., Hawkey, K., Ingram, M., Atkinson, A. and Smith, M. (2005) *Creating and Sustaining Effective Professional Learning Communities*. Nottingham: DfES Publications (Report number: RB637).

Bottery, M. (2006) Educational leaders in a globalising world: a new set of priorities? *School Leadership & Management*, 26(1), pp. 5–22.

Bryman, A. (2004) *Social Research Methods*. 2nd edition. Oxford: Oxford University Press.

Bubb, S. and Earley, P. (2008) *From self-evaluation to school improvement: The importance of effective staff development*. Reading: CfBT Research Report.

Bubb, S. and Earley, P. (2010) *Helping Staff Develop in Schools*. London: Sage Publications.

Butler, F.M. (1999) Reading Partners: Students Can Help Each Other Learn to Read. *Education and Treatment of Children*, 22(4), pp. 415–426.

Coldwell, M. and Simkins, T. (2011). Level models of continuing professional development evaluation: a grounded review and critique. *Professional Development in Education*, 37(1), pp.143–157.

Collins, T. and Dolan, R. (2011) Leadership and the Curriculum: A Question of Trust. In: O'Sullivan, H. and West-Burnham, J. (eds.) *Leading and Managing Schools*. London: Sage Publications, pp. 75–88.

- Centre for the Use of Research and Evidence in Education (CUREE) (2008) *Qualitative study of school-level strategies for teachers' CPD*. London: Cambridge University and The Open University.
- Darling-Hammond, L. (1997) *The Right to Learn: A Blueprint for Creating Schools that Work*. San Francisco: Jossey-Bass Publishers.
- Department of Education and Skills (DES) (2011) *Literacy and Numeracy for Learning and Life*. Dublin: DES.
- Desimone, L.M. (2009). Improving Impact Studies of Teachers' Professional Development: Toward Better Conceptualizations and Measures. *Educational Researcher*, 38(3), pp.181–199.
- Earley, P. and Bubb, S. (2004) *Leading and Managing Continual Professional Development*. London: Sage Publications.
- Earley, P. and Porritt, V. (2010) *Effective Practices in Continuing Professional Development*. London: Institute of Education University of London.
- Fallon, G. and Barnett, J. (2009) When is a Learning Community just a Pseudo Community? Towards the Development of a Notion of an Authentic Learning Community. *International Studies in Educational Administration*, 37(2), pp. 3–24.
- Frost, D. and Durrant, J. (2003) *Teacher-led Development Work*. London: David Fulton.
- Gray, D.E. (2004) *Doing Research in the Real World*. London, Thousand Oaks, New Delhi: Sage Publications Ltd.
- Guskey, T.R. (1996) Jointly Planning Staff Training. *School Administrator*, 53(11), pp. 33–34.
- Guskey, T.R. (2000) *Evaluating professional development*. Thousand Oaks, CA: Corwin.
- Guskey, T.R. (2002) Does it make a difference? Evaluating professional development. *Educational Leadership*, 59(6), pp. 45–51.
- Guskey, T.R. (2005) 'A Conversation with Thomas R. Guskey'. *The Evaluation Exchange*, 11(4), pp. 12–15. [Online.] Available from: <http://www.hfrp.org/evaluation/the-evaluation-exchange/issue-archive/professional-development/a-conversation-with-thomas-r.-guskey> (Accessed 23 March 2009).
- Hall, G. E. and Hord, S. M. (1987) *Change in schools: facilitating the process*. Albany: SUNY Press.
- Hammersley, M. (2007) *Educational Research and Evidence-based Practice*. London: Sage Publications Ltd.

Hodkinson, H. and Hodkinson, P. (2005) Improving Schoolteachers' Workplace Learning. *Research Papers in Education*, 20(2), pp. 109–131.

Holton, E. (1996) The flawed four-level evaluation model. *Human Resource Development Quarterly*, 7(1), pp.5–22.

Author and Co-author (2009) Peer Tutoring: An Inclusive Approach to Supporting Reading: Case study in five disadvantaged schools. *Journal of the Irish Learning Support Association*, 31, pp. 50–61.

Author (2011) The role of leadership in developing and sustaining teachers' professional learning. *Management in Education*, 25(4), pp. 149–155.

Author (2012) *Developing and Sustaining Teachers' Professional Learning: A Case Study of Collaborative Professional Development*. Unpublished EdD Educational Research and Development dissertation. University of Lincoln. [Online] Available from: (<http:xxxxxxxxxxxxxxxx>) (Accessed 18 November 2012).

Kirkpatrick, D. (1959) Techniques for evaluating training programs. *Journal of ASTD*, 11, pp.1–13.

Kitching, K., Morgan, M. and O'Leary, M. (2009) It's the little things: exploring the importance of commonplace events for early-career teachers' motivation. *Teachers & Teaching*, 15(1), pp. 43–58.

Lincoln, Y.S. and Guba, E.G. (1985) *Naturalistic Inquiry*. Newbury Park, CA: Sage Publications Ltd.

MacNeill, M., Cavanagh, R.F. and Silcox, S. (2005) Pedagogic Leadership: Refocusing on Learning and Teaching. *International Electric Journal for Leadership and Learning*, 9(2) [Online]. Available from University of Calgary at: <<http://www.ucalgary.ca/iejll>> (Accessed 14 January 2009).

Miles, M.B. and Huberman, A.M. (1994) *Qualitative Data Analysis: An Expanded Sourcebook*. 2nd edition. Thousand Oaks, CA: Sage.

Muijs, D. and 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), pp.195–211. [Online] Available at: <http://search.ebscohost.com/login.aspx?direct=true&db=tfh&AN=31474644&site=ehost-live>. (Accessed 3 December 2011).

Norris, N.D. (2004) Staff Development: Creating a Community of Learners. *International Electronic Journal for Leadership in Learning*, 8(2). [Online.] Available from University of Calgary at: <<http://www.ucalgary.ca/iejll>> (Accessed 12 January 2009).

NSW Institute of Teachers (2007) Continuing professional development policy. [Online.] Available from: <<http://www.nswteachers.nsw.edu.au/IgnitionSuite/uploads/docs/Continuing%20Prof>

[essional%20Development%20Development%20Policy.pdf](#)> (Accessed 24 November 2008).

Office for Standards in Education (Ofsted) (2006) *The Logical Chain*. London: Ofsted.

Opfer, V.D., Pedder, D. and Lavicza, Z. (2010) The role of teachers' orientation to learning in professional development and change: A national study of teachers in England. *Teaching and Teacher Education*. [Online.] Available from: <http://www.darleenopfer.com/File/TATE1506_proof.pdf> (Accessed 24 May 2011).

O'Sullivan, H. (2011) Leading and Managing Professional Learning in Schools. In: O'Sullivan, H. and West-Burnham, J. (eds.) *Leading and Managing Schools*. London: Sage Publications Ltd., pp. 111–125.

Pedder, D., Storey, A. and Opfer, V.D. (2008) *Schools and continuing professional development (CPD) in England: State of the Nation research project report for the Training and Development Agency for Schools*, University of Cambridge, UK. [Online.] Available from: <http://www.tda.gov.uk/cpd-leader/effective-cpd/~media/resources/cpd-leader/effective-cpd-research/cpd_statnation_report_survey.pdf> (Accessed 20 December 2010).

Priestley, M., Miller, K., Barrett, L. and Wallace, C. (2011) Teacher Learning Communities and Educational Change in Scotland: the Highland Experience. *British Educational Research Journal*, 37(2), pp. 265–284.

Prosser, M. and Trigwell, K. (1999). *Understanding Learning and Teaching: The Experience in Higher Education*, Philadelphia USA: Open University Press.

Punch, K.F. (2009) *Introduction to Research Methods in Education*. Thousand Oaks CA: Sage Publications Ltd.

O'Sullivan, H. (2011) Leading and Managing Professional Learning in Schools. In: O'Sullivan, H. and West-Burnham, J. (eds.) *Leading and Managing Schools*. London: Sage Publications Ltd., pp. 111–125.

Rhodes, C., Stokes, M. and Hampton, G. (2004) *A Practical Guide to Mentoring, Coaching, and Peer-Networking: Teacher Professional Development in Schools and Colleges*. New York: RoutledgeFalmer.

Robson, C. (2002) *Real World Research*. 2nd edition. Oxford: Blackwell.

Silverman, D. (2000) *Doing Qualitative Research: A Practical Handbook*. London: Sage Publications.

Smith, R.G. (2007) Developing professional identities and knowledge: becoming primary teachers. *Teachers & Teaching*, 13(4), pp. 377–397.

Stake, R.E. (1967) The countenance of educational evaluation. *Teachers College Record*, 68(7), pp. 523 Stevenson, H. (2008) Teacher Development through

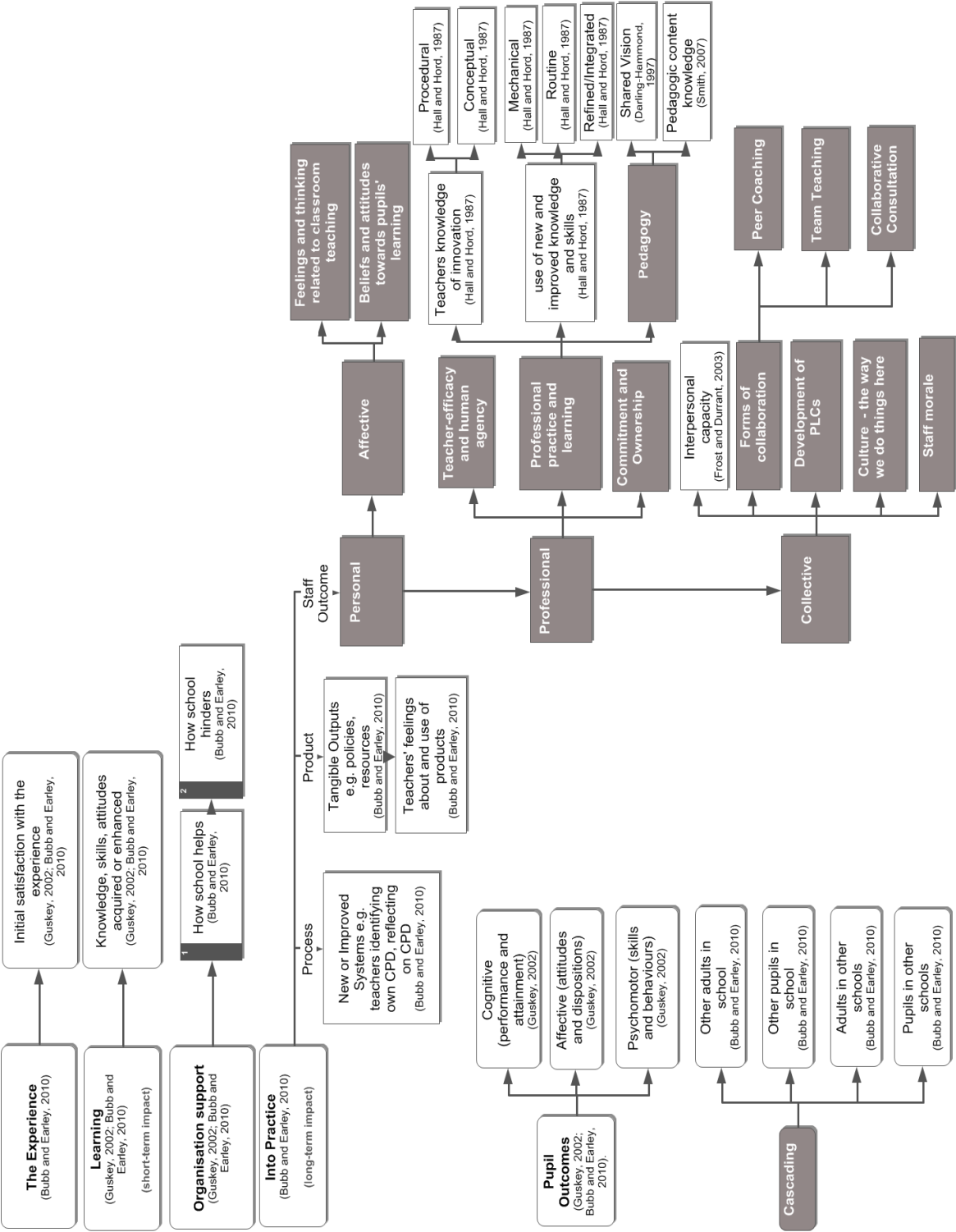
Practitioner Research and Cross-school Collaboration: Possibilities and Limitations.
In: Ling-po Shiu and Chi-Kin Lee, J. (eds.) *Developing Teachers and Developing
Schools in Changing Contexts*. The Chinese University Hong Kong: The Chinese
University Press, pp. 337–358.

Stoll, L. and Fink, D. (1996) *Changing Our Schools: Linking School Effectiveness
and School Improvement*. Maidenhead: Open University Press.

Topping, K. (1988) *The Peer Tutoring Handbook*. London: Croom Helm.

Yin, R.K. (1994) *Case Study Research: Design and Methods*. 2nd edition. Thousand
Oaks, CA: Sage Publications Ltd.

Yin, R.K. (2009) *Case Study Research Design and Methods*. 4th edition. Thousand
Oaks, CA: Sage Publications Ltd.



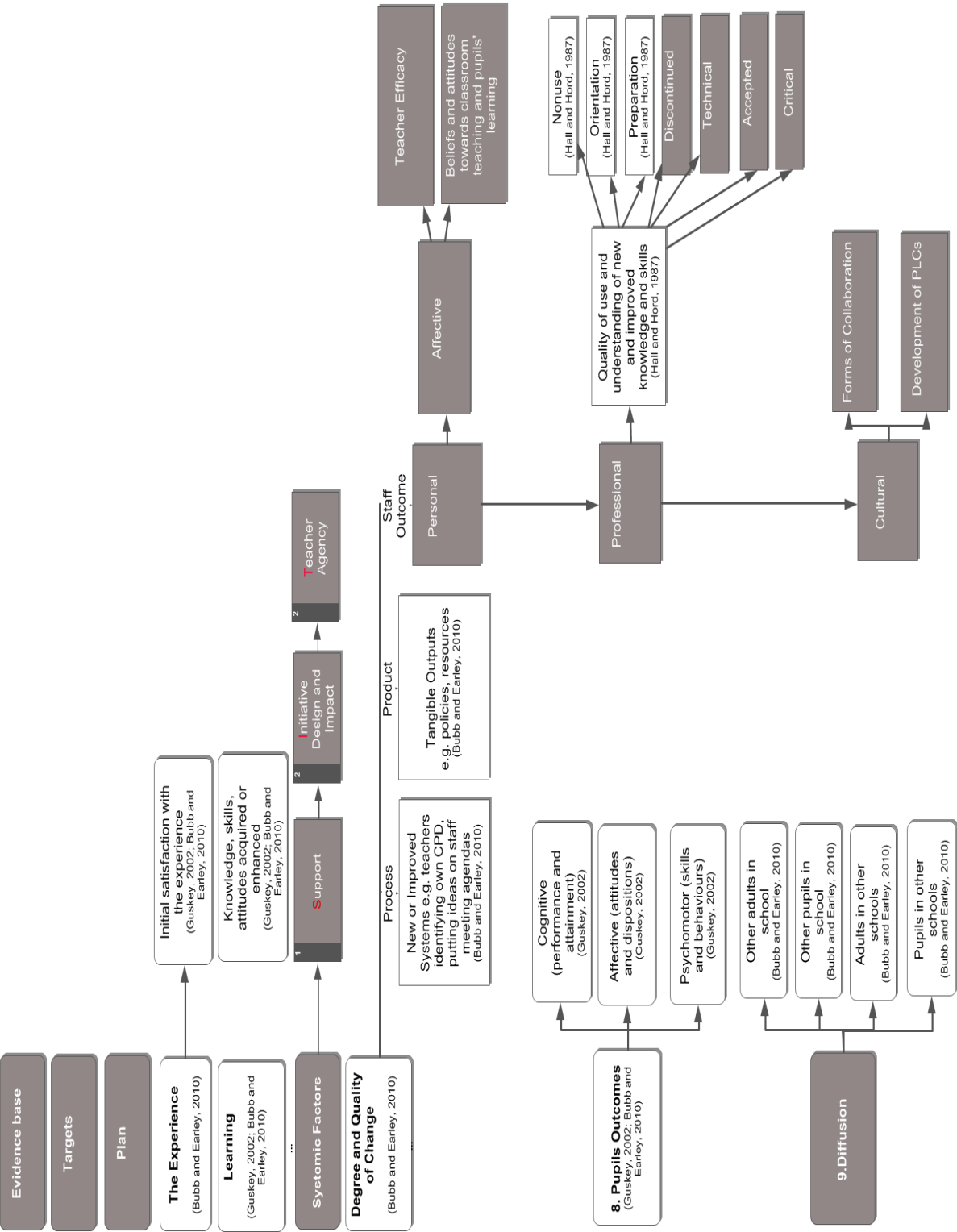


Table 1: Comparison of evaluation models

Guskey (2002)	Bubb and Earley (2010)
<ol style="list-style-type: none"> 1. Participants' reactions 2. Participants' learning 3. Organisation support and change 4. Participants' use of new knowledge and skills 5. Students' learning outcomes 	<ol style="list-style-type: none"> 1. Baseline picture 2. Goal 3. Plan 4. The experience 5. Learning 6. Organisational support 7. Into Practice- Degree and quality of change (process, product or staff outcome) 8. Students' learning outcomes 9. Other adults in school 10. Other students 11. Adults in other schools 12. Students in other schools

Table 2 Levels of Use of the Innovation

	Level of Use	Behaviours
Users	Renewal	Makes major modifications in the innovation or their innovation use to improve the impact on pupils Examines new developments in the field
	Integration	Commitment to use the innovation with other teachers to provide a collective impact on pupils Makes changes to accommodate the use of innovation with another teacher
	Refinement	Makes changes to enhance the impact on pupils in their class
	Routine	Established use of the innovation Little thought about improving innovation use Not making any changes to it
	Mechanical	Concerned with logistics and organisational issues Clings to the user guide Makes changes to suit user needs Focuses on short-term, day-to-day use of the innovation
Nonusers	Preparation	Makes a decision to use the innovation Preparation and planning for the first use of the innovation
	Orientation	Takes action to learn more detailed information about the innovation e.g. Looks for information about the innovation- talking to others, attending a workshop Explores the possibilities for use of the innovation No commitment to use the innovation
	Nonuse	Absence of innovation-related behaviour - no knowledge, involvement and doing nothing toward becoming involved