Australian Council for Educational Research
ACEReSearch

Professional Development for Teachers and School Leaders | Teaching and Learning and Leadership

1-1-2003

Evaluating the quality and impact of professional development programs

Lawrence Ingvarson
ACER, Lawrence.Ingvarson@acer.edu.au

Marion Meiers
ACER, Marion.Meiers@acer.edu.au

Adrian Beavis
ACER, adrian.beavis@acer.edu.au

Follow this and additional works at: http://research.acer.edu.au/professional_dev

Part of the Educational Assessment, Evaluation, and Research Commons

Recommended Citation
Ingvarson, Lawrence; Meiers, Marion; and Beavis, Adrian, "Evaluating the quality and impact of professional development programs" (2003).
http://research.acer.edu.au/professional_dev/3

This Report is brought to you for free and open access by the Teaching and Learning and Leadership at ACEReSearch. It has been accepted for inclusion in Professional Development for Teachers and School Leaders by an authorized administrator of ACEReSearch. For more information, please contact mcdowell@acer.edu.au.

Evaluating the quality and impact of professional development programs

Lawrence Ingvarson, Marion Meiers and Adrian Beavis

Teaching and Learning Research Program, Australian Council for Educational Research

Professional development for teachers is now recognised as a vital component of policies to enhance the quality of teaching and learning in our schools. Consequently, there is increased interest in research that identifies features of effective professional learning. Considerable funds are allocated to a wide variety of professional development programs from a variety of sources. As investment increases, policy makers are increasingly asking for evidence about its effects not only on classroom practice, but on student learning outcomes. They are also looking for research that can guide them in designing programs that are more likely to lead to significant and sustained improvement in student opportunities to learn.

There is a need, therefore, for more sophisticated methods of evaluating professional development, with the capacity to meet these information needs. In the not too distant past, when many professional development courses placed teachers in the role of an audience, questionnaires distributed at the door as teachers left sufficed. Strategies for professional development have now become much more complex, long term and embedded in schools. Major funds may be allocated to training school-based staff developers and providing them with time release, developing curriculum support materials, time release, on-line learning and so on.

The kinds of questions that evaluators now need to answer are much more penetrating than questions such as "What did you learn from the workshop?" They are questions about program logic and the presumed links between professional learning strategies, and changes in teacher knowledge, classroom practices and student outcomes. These questions call for large-scale studies with the capacity to test these relationships across large numbers of different professional development programs.

Purpose of this paper

The purpose of this paper is to review recent work that ACER has been doing to improve the usefulness of evaluations of professional development programs. This work includes the development of research-based instruments to measure:

- the nature and quality of the *processes* used to promote teacher learning;
- the *impact* of programs on teacher knowledge, practice and student outcomes;
- the relationships between these process and impact measures.

The paper is based on approaches developed as part of four recent evaluations of professional development programs. These include:

- three evaluations of the Commonwealth Government's Quality Teacher Program, as implemented in three separate states: New South Wales, CEC Victoria and the Northern Territory; and
- a major research study funded by the Commonwealth Government investigating the links between professional development and student learning outcomes.

Key features of the ACER approach to evaluation

Cross-program analysis

In each of these evaluation studies, data was gathered from a number of PD programs. In evaluating the NSW QTP, for example, data was gathered from 41 programs and 1731 teachers. In conducting all four evaluations, data was gathered from a total of 3250 teachers who had participated in eighty different professional programs across all states in Australia. These studies provided a unique opportunity to conduct research looking at the differential impact of a wide range of PD strategies.

Participants in each of these programs were invited to complete a common survey instrument, which asked them to describe both the processes of learning that they had experienced and the impact of these programs on their knowledge, practice, sense of efficacy, and their students' learning. The survey also asked participants about the impact of the programs on the nature and extent of collaborative work amongst colleagues in their schools. The extent to which programs strengthened, or integrated with professional community activity was a significant predictor of impact.

As might be expected, there were significant differences between programs in the mode of delivery and in the extent to which teachers reported that programs had influenced their practice and benefited their students. These differences opened up the

possibility for cross-program analyses that might:

a) increase understanding of those features of project design and delivery that might explain variation in impact;
b) identify school level factors that influence or mediate the outcome of the projects.

Another feature of these studies was that teachers were surveyed at least three months after participating in a program, which provided them with the opportunity to gauge the impact of programs on their practice. Unfortunately this delay was at a some cost to response rates to our mailed surveys, which varied, but averaged around 50%.

Research-based conceptual framework

These analyses called for the development of a conceptual framework to guide the evaluation. The ACER approach to evaluation in each of the four studies was based on the theoretical framework, shown in Figure 1. It presents a model of the main program features that might explain variation in the reported impact of PD programs. The framework was based on a review of recent research into the characteristics of effective professional development programs (Kennedy, 1998; Wilson & Berne, 1998; Garet et al., 2001; Sykes, 2002; Ingvarson & Meiers, 2003; Cohen & Hill, 2000, Hawley & Valli, 1999; Guskey, 2002; Loucks-Horsley et al. 1998; Supovitz, 2001). This research has become increasingly sophisticated over recent years. (Ingvarson, 2002) and provides a firmer foundation on which to develop models to account for the relative differences in the effectiveness of professional development programs.

Figure 1 distinguishes four, linked, types of impact resulting from PD programs. These include impact on teachers' knowledge and practice, student learning and teacher efficacy. The model also includes background (control) variables, structural features, such as the duration of the program and opportunity to learn features, such as "active learning", or "follow up". (Details of how these variables are measured are provided below.)

Mediating variables

Many PD programs aim to strengthen professional community in schools in order to enhance the impact of their programs on classroom practice. Therefore, professional community is included in our model as a mediating variable. In measuring professional community teachers are asked to respond to items such as:

- Teachers at my school discuss teaching and learning more with their colleagues
- Teachers have increased their collaboration in planning, teaching and assessment activities
- I have passed ideas I learned from the project on to other teachers in my school

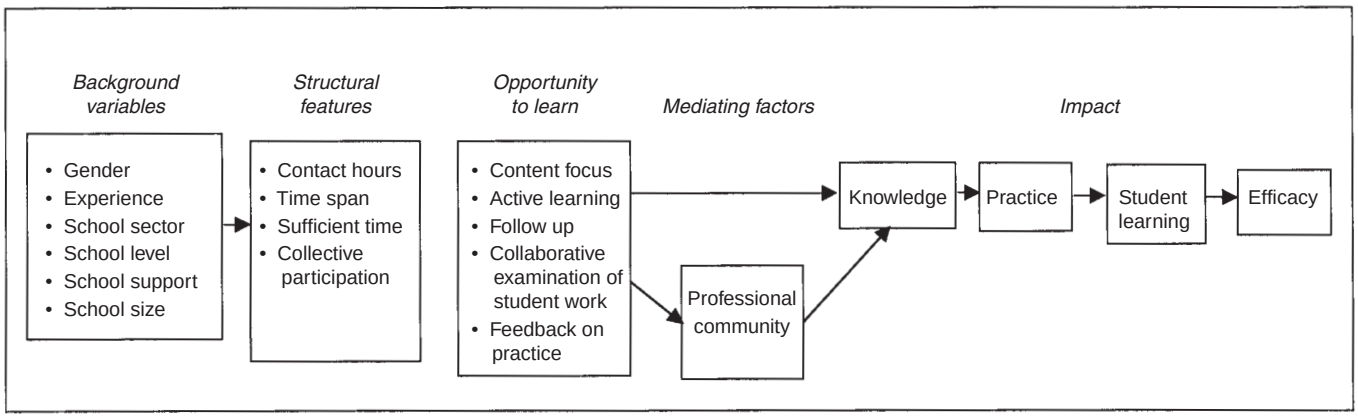
Analyses of program logic and theory of action

The first step in any evaluation is to clarify the focus of the evaluation; that is, to define exactly what it is that is to be evaluated. This involves identifying the key design and process features of the approach being used in a professional development program – what the program looks like in practice and how it is meant to work. This task is not always as straightforward as it may seem, as program designers may not have articulated these matters before.

A feature of the ACER approach to evaluation is the emphasis placed on working in close collaboration with policy makers and providers to identify the essential and critical features of the professional development model they are using. This includes identifying the assumptions about teacher learning on which their models are based, and teasing out the theory of action underlying their programs (how the features of the proposed model link to each other and how they will lead to change).

In working with program designers, ACER staff draw extensively from recent research on the critical features of effective professional development programs (Hawley & Valli, 1999; Ingvarson & Meiers, 2003).

Figure 1 Relationships between structure, learning processes and impact of professional development programs



Use is made of other researchers (e.g. Loucks-Horsley et al., 1998; Sykes, 2002) who provide useful guides to the major types of strategies used to promote professional learning. Heller et al. (2003) and Killion (2003) provide approaches that help to identify the logic underlying programs and how the pieces fit together to promote effective teacher learning.

The outcome of this collaborative work with designers is the production of program profiles that identify:

- the main components in the design (inputs, structures, activities, initial outcomes, longer term outcomes);
- how these components are expected to link together in practice to promote teacher learning; and
- the theory of action that underpins their project (i.e. the mechanisms by which project activities will lead to change in classroom practice).

These program profiles help to provide a clearer idea of the kind of data that needs to be collected in conducting evaluations with the capacity to test the assumptions underlying the professional learning models and provide useful information for refining the programs.

Measures of professional development strategies and learning processes (opportunity to learn)

While we use project profiles to clarify what is to be evaluated, we have found that we can not rely on them entirely as accurate measures of teachers' actual opportunities to learn during programs. A special problem in conducting evaluations of professional development programs is gathering data about what teachers actually do and how they learn in the program; what roles they play as learners and the nature and extent of their actual opportunities for learning.

Designers of professional development programs select from a wide range of strategies to promote professional learning. They often describe the strategies they have chosen in ways that are not particularly helpful for research purposes. They may use terms such as, 'hands on', 'action research', 'workshops', 'training sessions', 'case methods'. What these terms actually mean in terms of teacher learning processes is not always clear. To make the research task even more complex, designers often say they use a large number of these strategies in their programs. So we found it difficult to gain useful measures of actual teacher learning processes by asking program designers about the strategies that characterise their programs.

Rather than relying on what the providers say about the design features and learning processes of their programs, we prefer to rely on what teacher-participants report about their experience in the program – their actual opportunities to learn. A program may be advertised as 'action research', for example, but teachers' actual experience may be quite different. Program designers may claim to have provided follow up support, but teachers may not have received it.

As indicated above, research now provides a firmer foundation on which to develop models that might account for variation in the effectiveness of professional development programs. The evaluation team used this research to create an instrument for measuring the quality of opportunities for teachers to learn. In developing this instrument (*The Quality of Professional Learning Index*) we used our review of the research literature to identify a number of characteristics of effective professional development. These included:

- content focus
- follow up
- active learning
- feedback
- collaborative examination of student work.

Each of these measures is described briefly below (it is important to note that this instrument is being refined continually in the light of research).

Content focus

Recent research (Kennedy, 1998) indicates the importance of *what* teachers have the opportunity to learn during professional development programs – this research indicates that the substance of what teachers learn is more important than the form or structure of the program (e.g. whether programs are school-based or not, collaboratively planned or not, extended over time, etc.). In summary, this research indicates that professional learning is more likely to improve student learning outcomes if it increases teachers' understanding of the content they teach, how students learn that content and how to represent and convey that content in meaningful ways (Cohen & Hill, 2000).

To measure content focus, teachers are asked about the emphasis given to four aspects of content: content or subject knowledge, knowledge of how students learn content, knowledge of methods of teaching content and models to illustrate those methods of teaching of that content.[1]

Active learning

Recent research confirms the importance of importance of teachers being actively engaged in their own learning, but it is the nature of this engagement that seems to matter as much, if not more, than the level.

[1]To measure *content focus*, an index was developed based on four items. Teachers responded to these items on a four-point scale from '1 = no emphasis' to '4 = major emphasis'. The scores of each of these items were averaged to give a measure of content focus. A similar process was used to construct all measures

Effective professional development programs draw teachers into an analysis of their current practice in relation to professional standards for good practice. They also draw teachers into close comparison of what their students are learning in relation to what students of that age and circumstance are capable of learning.

To measure active learning, teachers are asked about the extent to which a program engaged them actively in reflecting on their practice, in identifying specific areas of their practice that they needed to develop, and gave them opportunities to test new teaching practices.

Feedback

Feedback on practice has long been recognised as a vital requirement for professional development programs that aim to help teachers develop new skills and integrate them into their practice (Joyce & Showers, 1982). Effective integration of new skills requires programs to have a clear theoretical foundation supported by research, modelling in real settings, and opportunities to practice the new skills and receive feedback from a coach or supporting teacher. Most of the programs we have evaluated recently aimed to help teachers learn new skills. However, we found that few participants actually received assistance and feedback in their classrooms during the critical and difficult implementation phase when they were trying out new practices.

To measure feedback, teachers are asked about the number of times they received feedback on their teaching from other teachers or people involved in the program; and the number of times their teaching was observed by others involved in the program (e.g. from a mentor, or in a team teaching situation).

Collaborative examination of student work

Effective professional development programs lead teachers to examine their students' work in relation to external reference points or standards. Hawley and Valli's (1999) review of research rates this feature as a critical component of effective professional learning programs. It has become clear over recent years that teachers gain a great deal of valuable learning from opportunities to examine student work in collaboration with colleagues - especially their own students' work, and in relation to standards for what students should know and be able to do. Collaborative analyses of student work opens up many avenues for teachers to de-privatise their practice and learn from each other. It also leads to deeper understanding of student learning outcomes and greater discrimination about what counts as meeting those objectives.

To measure *collaborative examination of student work*, we developed an index based on the extent to which teachers said they received opportunities to collaborate with colleagues in examining their own students' work as well as that of other teachers.

Follow up

Follow up support to teachers during the implementation phase of change has long been identified as an important feature of more effective programs (Fullan, 1982). Perhaps the strongest criticism of many professional development programs over the years has been the lack of built in provision for 'at the elbow' support for teachers in their classrooms as they apply new ideas and skills (Huberman & Miles, 1984).

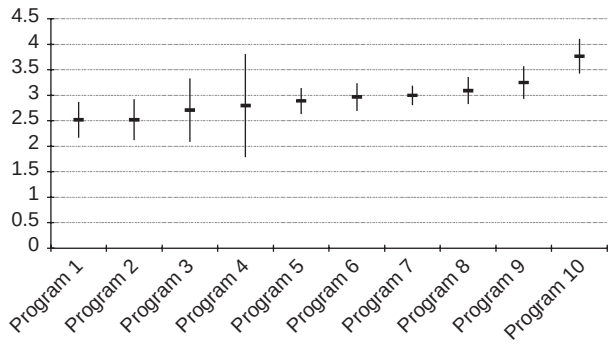
To measure *follow-up* we developed an index based on the extent to which teachers reported that a program provided time for follow-up and ongoing assistance in their school or classroom to help them implement changes advocated in the program and opportunities to practice their new learning

Factor analysis confirmed the scales used to measure the five opportunity to learn constructs described above. Details about the psychometric properties of these opportunity to learn variables will be provided in a fuller version of this paper.

Presenting findings about opportunity to learn

Figure 2 shows an example of how we can present findings about these opportunity to learn measures; in this case the level of *content focus*. These findings come from an evaluation of ten professional development programs across Australia (Programs 1-10). Figure 2 shows, for example, significant variation across the ten programs in terms of our measure of content focus. Teachers in Programs 8, 9 and 10 reported that these programs placed more emphasis on content than Program 1, 2 and 3.

Figure 2
The extent to which the program was reported to have a focus on content, showing mean emphasis level and 95% confidence intervals



Measures of impact based on teaching standards

Another feature of the ACER approach to evaluating professional development programs is the method

developed for measuring impact. In order to conduct research based on the conceptual model in Figure 1, it was necessary to develop a common framework of measures for assessing impact. The ACER evaluation team created a new way of conceptualising and identifying outcomes of PD programs based on standards for effective teaching (Ingvarson, 1998; 2002). We argued that the quality of impact of a PD program should be measured primarily, not in terms of whether it met the developers' objectives, but in terms of the extent to which the program moved teachers' practices towards those associated with research-based standards for effective teaching (Ingvarson, 1998; 2002). (These objectives may be the same, but not necessarily.)

We developed four aspects of *impact* for our recent evaluations: impact on teachers' knowledge; impact on teachers' practice; impact on student learning outcomes; and, impact on teacher efficacy. Teachers report their responses to the following items on a four-point scale from strongly agree to strongly disagree.

Knowledge

Teachers are asked to indicate the extent to which their participation in the PD program has led to increased knowledge of: the content they teach, teaching and learning strategies appropriate to the content they teach, how students learn the content, individual differences amongst students and how to cater for their needs, how to link assessment into the teaching and learning cycle, classroom organisation and management, materials and resources available in their area of teaching.

Practice

Teachers are asked whether, as a result of their participation in the PD program, they now:

- make clearer links between their teaching goals and classroom activities;
- manage classroom structures and activities more effectively;
- use more effective teaching and learning strategies appropriate to the content that they teach;
- use more effective teaching and learning strategies appropriate to the classroom context;
- use teaching and learning strategies that are more challenging and engaging;
- are better able to meet the individual learning needs of their students;
- link assessment into the teaching and learning cycle more effectively;
- provide more effective feedback to their students to support their learning;
- engage students in higher order thinking;
- access and use materials and resources more effectively.

Student learning outcomes

Teachers are asked whether, as a result of the PD program, their students now:

- have fewer difficulties in understanding what they are being taught;
- are learning more purposefully;
- are more actively engaged in learning activities;
- demonstrate enhanced learning outcomes;
- access and use materials and resources more effectively.

Teacher efficacy

Teachers are asked about the extent to which they agree or disagree with the following statements:

As a result of the PD program:

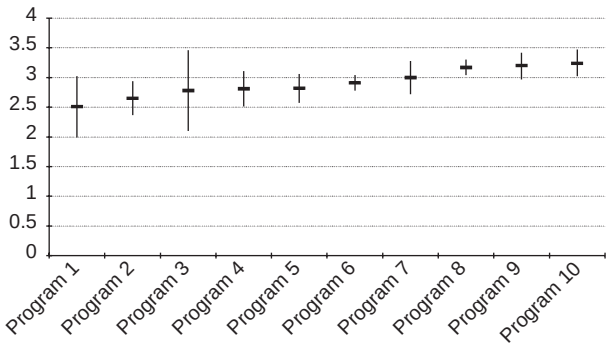
- My ability to meet the learning needs of my students has been expanded
- My confidence as a teacher has increased

All the above measures had strong scale characteristics and they proved to be sensitive to differences across programs

Comparisons of PD programs in terms of impact

The above measures of impact enabled comparisons to be made across PD programs, such as illustrated in Figure 3 below for impact on practice. Figure 3 compares ten major PD programs. Figure 3 shows that Programs 1 and 2 programs had statistically lower average levels of reported impact on practice than Programs 9 and 10.

Figure 3
The extent to which teachers agree that there was a positive change in their professional practice, showing mean agreement and 95% confidence intervals



Findings

Space here precludes the presentation of anything more than a sample of the types of analyses undertaken in these studies and the findings.

Regression analysis

Blockwise regression analysis is usually conducted in analysing relationships between components of the conceptual model in Figure 1 above. This procedure is based upon a least-squares algorithm to estimate the strength of the linear relationship between the dependent variable and a set of independent variables. Results from the ACER evaluation of the CEC Quality Teacher Program in Victoria are summarised in Table 1. The order in which these variables are entered into the equation is determined by the theory underlying the research (as summarised in Figure 1). There were six control, or background (exogenous), variables in this model, and three blocks of intervening (endogenous) variables: structural feature, learning process and professional community.

Table 1 shows the standardised regression coefficients and significance levels for each of the predictors in the model. The use of standardised co-efficients permits easy comparison of the strength of associations within the model. For example, a standardised beta coefficient of 0.27 is three times as strong in its effect as one of 0.09. When examining these effects it is important to remember that they are net of the effects of other variables in the model. The regression analysis thus shows the unique contribution that each variable makes to changes in the dependent variable. Table 1 (below) also shows the proportion of variance explained by the models (R^2).

The full model accounted for around 59% of the variance in the dependent variable (reported changes in teaching practice) – which means that several features in our model are reasonably good predictors of whether teachers rate professional development programs as effective in terms of changing practice.

The main message from Table 1 (and from other ACER evaluations of professional development programs) is that the block of variables associated with opportunity to learn has significant effects on our measures of impact. The block of variables seen in Table 1 from *Content Focus* to *Feedback*, together contribute importantly to predicting levels of reported changes in teacher knowledge, practice and teacher efficacy.

Table 1 also indicates that the background variables (non-project related) have weak links to impact. However, the level of associated professional community activity generated by a program, as a mediating variable, has a significant effects on teacher knowledge and practice. Improved practices and improved student learning outcomes, not surprisingly, are strongly associated with teacher reports about the impact of programs on their efficacy.

Similar results could be presented across the four major evaluation studies listed above that ACER has completed recently. A later paper will provide a much more extensive analysis of the findings and a discussion of implications for future investment in professional development for teachers.

Limitations

The approach to evaluation described above is based primarily on teacher self-report data. Given the time frame and the level of resources usually allocated to evaluations of professional development programs, there is often little opportunity to gather first-hand evidence about changes in teacher knowledge, practice, efficacy and students' learning outcomes. However, recent studies (e.g. Mayer, 2001) indicate that it is reasonable to place a certain level of

Table 1 Relationship between background variables, structural features, opportunity to learn, professional community in the school, and teacher knowledge, teacher practice, student learning and teacher efficacy

	Content	Active	Follow up	Collaboration	Feedback	Professional Community	Knowledge	Practice	Student Outcomes	Teacher Efficacy
Gender (F = 0 M=1)	0.01	-0.02	**0.08**	0.00	0.03	0.01	-0.01	-0.01	-0.02	0.04
School size	-0.04	-0.01	-0.01	-0.04	0.00	-0.03	-0.04	0.01	-0.07	-0.01
Primary = 1 else = 0	**0.16**	-0.02	**0.10**	0.00	-0.06	0.02	**0.09**	-0.02	**0.13**	-0.05
Vounteered=1 Directed=0	0.01	**0.06**	-0.01	**0.06**	-0.01	0.00	0.04	0.05	**-0.06**	0.02
School Group 1=Y 0=No	0.06	0.02	**0.09**	**0.15**	0.03	**0.14**	0.01	-0.03	0.05	0.02
PD Support in School	**0.14**	0.05	**0.12**	0.03	0.06	**0.13**	-0.05	-0.01	0.05	-0.03
Duration Hours	**0.21**	**0.12**	**0.07**	0.04	0.05	0.01	0.05	-0.05	**0.07**	-0.02
Span-months	**0.11**	0.02	0.03	**0.13**	0.05	-0.01	-0.04	0.07	-0.03	0.05
Sufficient Time	**0.19**	**0.11**	**0.09**	0.02	0.04	0.03	0.06	0.01	0.04	**0.06**
Content		**0.55**	**0.08**	**0.28**	0.07	**0.10**	**0.35**	**0.11**	**0.14**	**0.11**
Active			**0.47**	**0.14**	0.00	0.00	**0.09**	**0.09**	-0.02	**0.22**
Follow up				**0.30**	**0.23**	**0.25**	**0.09**	**0.10**	**0.10**	0.00
Collaboration					**0.20**	**0.21**	0.07	**0.11**	0.03	**0.10**
Feedback						**0.16**	0.04	0.02	0.06	**0.09**
Professional Community							**0.12**	**0.15**	**0.11**	**0.08**
Knowledge								**0.57**	0.04	0.02
Practice									**0.55**	**0.17**
Student Outcomes										**0.31**
R-square Adjusted	0.20	0.42	0.43	0.49	0.15	0.39	0.38	0.59	0.51	0.50

confidence in surveys that rely on teachers' reports about their practice. Reliability of these self-report data increases with more specific measures, as used in the ACER approach. Also, teachers are not reluctant to speak their minds frankly when it comes to assessing the value of professional development programs. There is little reason to think that their responses might be biased one way or another, or any desire to please, especially when, in studies such as the above, they are contacted several months at least after the programs have finished.

References

Cohen, D K and Hill H C, (2000) "Instructional policy and classroom performance: the mathematics reform in California", *Teachers College Record 102*, 2, February 2000

Fullan, M (1982) *The New Meaning of Educational Change*, (3rd edition) Teachers' College Press, New York

Garet, M S, Porter, A, Desimone, L, Birman, B F, Yoon, K S (2001) What Makes Professional Development Effective? Results From a National Sample of Teachers, in *American Educational Research Journal, Winter 2001, vol 38*, pp 915-945

Guskey, T.R. & Sparks, D (2002)"Linking Professional Development to Improvements in Student Learning", paper presented at the annual meeting of the American Educational Research Association, April 2002

Hawley and Valli, (1999) "The essentials of effective professional development: A new consensus" in Darling-Hammond, Linda and Sykes, Gary, Editors (1999) *Teaching as the Learning Profession. Handbook of Policy and Practice*, Jossey-Bass Publishers, San Francisco, 1999

Heller, J.I., Daehler, K.R., & Shinohara, M. (2003). *Connecting all the pieces*. Journal of Staff Development. Vol. 24(4), 36-41.

Huberman, M., & Miles, M. (1984). *Innovation up close*. New York: Plenum Press.

Ingvarson, L.C. (1998). Teaching standards: foundations for the reform of professional development. In A. Hargreaves, A. Lieberman, M. Fullan & D. Hopkins (Eds). *International handbook of educational change.* Dordrecht: Kluwer.

Ingvarson, L.C. (2002). *Building a Learning Profession*. Paper No 1, Commissioned Research Series, Australian College of Education. Canberra: Australian College of Education.

Ingvarson, L and Meiers, M, (In preparation) "Investigating the links between teacher professional development and student learning outcomes." Report for DEST, Canberra.

Joyce, B., & Showers, B. (1995). *Student achievement through staff development: Fundamentals of school renewal*. Second Edition. White Plains, NY: Longman.

Kennedy, M. (1998). *Form and substance in in-service teacher education* (Research monograph no. 13). Arlington, VA: National Science Foundation.

Killion, J. (2003). Solid footwork makes evaluation of staff development a song. Journal of Staff Development. Vol24(4), 14-21.

Loucks-Horsley, S; Hewson, P; Love; Stiles, K, The National Institute for Science Education, Corwin Press, (1998) "Strategies for Professional Learning" in *Designing Professional Development for Teachers of Science and Mathematics*

Supovitz, J. (2001). Translating teaching practice into improved student achievement. In S. Fuhrman (Ed.). *From the capitol to the classroom. Standards-based reforms in the states. The one hundredth yearbook of the National Society for the Study of Education, Part Two* (pp. 81-98). Chicago: University of Chicago Press.

Sykes, G. (2002). Professional Development for Teachers: Principles, Practices and Contexts. Paper prepared for the Learning First Alliance (Draft).

Wilson, S. and Berne, J. (1999) "Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development', in *Review of Research in Education*, Vol 24, Iran-Nejad and PD Pearson (eds) American Educational Research Association, Washington DC, pp 173-209