Bamber, Veronica (2008) *Evaluating lecturer development programmes: received wisdom or self-knowledge?* International Journal for Academic Development, 13 (2). pp. 107-116. ISSN 1360-144X

Evaluating Lecturer Development Programmes - Received Wisdom or Self-Knowledge? ***International Journal of Academic Development.*** **Vol 13, No 2, June 2008. Pp 107-116**

Evaluating Lecturer Development Programmes – Received Wisdom or Self-Knowledge?

Abstract

This paper argues that lecturer development programmes can best be evaluated by theory-informed, contextualised evaluations involving a structured approach. The self-knowledge obtained by course teams who engage with such an approach outweighs wisdom received from external evaluators within large-scale evaluation initiatives, although these can provide a useful backdrop. Since lecturer development programmes are change initiatives, the theory of change is an important starting point. The paper considers published evaluations of programmes, and outlines the evaluation processes in one case institution. A structured approach, using an evaluation framework, is suggested. Although many of the examples are UK-based, the approaches described will be relevant more broadly, and lessons for evaluators are outlined.

Introduction

This paper addresses three related research questions, regarding what can be learned from large-scale evaluations of lecturer development programmes (LDPs), what can be learned from published, smaller scale evaluations, and what parallels exist between published evaluations and the experience of a particular case institution.

Until recently, there was little published evaluation of LDPs, and little research on their impact (Coffey & Gibbs, 2000, p. 385). In fact, educational development activities generally suffer from a lack of systematic evaluation (Gosling, 2008, p. 38). A survey of 93 UK higher education institutions (Bamber, 2002) found that evidence of the impact of LDPs was usually anecdotal. The danger is that provision cannot be justified or defended, beyond it '*seems* like a good thing'. Perhaps more importantly, the existing literature on links between learning and teaching (e.g. Biggs (1999); Martin (1999); McAlpine (2000); Ramsden (1992); Trigwell, Prosser & Waterhouse (1999)) could be enriched if educational developers took a research-informed approach to evaluating their programmes, and published their findings. In the absence of such approaches, educational development may, indeed, be the 'precarious business' described by Gibbs and Coffey (2000, p. 39). This paper seeks to fill a gap by comparing the evaluations described in a small range of published papers. The approach taken was to scan the academic development literature for papers and key texts on LDP evaluation. The key texts included journal articles (eg from *Assessment and Evaluation in Higher Education, Teacher Development, Evaluation, Innovation in Education and Teaching*, and *IJAD* itself). Key words used were 'evaluation', 'higher education' and 'lecturer development'. Also useful were national reports, such as the Higher Education Academy *Formative Evaluation of Accredited Programmes* (Prosser et al, 2006), conference papers (eg from *ICED 1998)*, and a number of relevant book chapters. The analysis focused on the what, how, why (i.e. theory of change) and lessons of each evaluation approach.

All of the consulted texts point to the fact that LDP evaluation is far more complex than the commonly used 'happy sheet' form of workshop feedback might suggest, since there is no straightforward link between a change initiative and its outcomes (Trowler & Bamber, 2005, p. 88). It is notoriously difficult to find meaningful measures for improvements in lecturer development: the concept of 'measurement' may not even be appropriate. This paper suggests that developers need to gather and interpret a broad spectrum of data within their local context - a long-term, complex process, avoiding simplistic attempts to directly link professional development and student learning. Complexity entails uncertainty about causal effects, and requires a rich picture of lecturer experiences to judge the value of the LDP. To what extent can this picture be illuminated by larger, multi-institution evaluations found in published papers?

The value of large-scale evaluations

In their ground-breaking project, Gibbs and Coffey (2000) gathered data from trainee lecturers and their students in 22 universities in eight countries, following a research-informed methodology. This evaluation was underpinned by a theory of change that good teaching can positively affect student outcomes, that higher education (HE) teachers can be helped to improve the quality of their teaching, and that student-focused teachers contribute to better student learning. To test the theory, they administered questionnaires, the Approaches to Teaching Inventory (Trigwell and Prosser, 1999) and Teaching Methods Inventory (Coffey and Gibbs, 2002), on the lecturers' entry into and exit from the LDPs. Since Gibbs and Coffey (2000) sought correlations between lecturer development and student learning, the lecturers' students were also surveyed, using the Student Educational Experience Questionnaire (Marsh, 1982) and the Module Experience Questionnaire (Gibbs & Coffey, 2004). Although Gibbs and Coffey were unsure about the full methodological validity of the research, their approach demonstrated how a theory of change, theoretical underpinning and empirical data can be married together.

The UK Higher Education Academy (HEA) also undertook a multi-institution project, evaluating programmes in 32 institutions (Prosser et al, 2006). The approach was both quantitative (online questionnaire and the Approaches to Teaching Inventory (ATI - Trigwell and Prosser, 1999) and qualitative (focus groups with course participants and senior staff), and was underpinned by the same theory of change as the Coffey and Gibbs study. Despite some questionable methodological techniques, such as asking programme participants to complete the ATI as they remembered their ideas at the start of the programme (usually 2 years before), the evaluation went some way to filling a gap, providing comparative data across institutions, and individualised data for the universities involved. For example, broad findings about the effect of institutional, subject discipline and gender differences on individuals' perceptions of their LDP could be valuable when framing institution-specific evaluations.

However, a note of caution must be struck. Just as LDPs have a theory of change, so do evaluations. This means that each evaluation will take a particular philosophical and theoretical position. While the theory of change underpinning the HEA and Gibbs and Coffey evaluations suggested that student-focused teaching would bring deeper learning, Knight's (2006) evaluation started from a different theoretical and philosophical position: his large-scale evaluation project (questionnaire data from 171 respondents and a sample of interviews) focused on the nature of professional learning. Knight's theory of change stemmed from the writings of a different school from the previously mentioned studies, with the major message that LDPs did not lend themselves to professional learning. He cited, for instance, Blackler (1995), suggesting that, if professional knowledge is largely implicit, and implicit knowledge tends to be acquired unintentionally while doing a job, then learning to teach in HE is best done in the workplace, not in formal courses. A caution is then, that, while evaluation should be informed by the literature, it needs to be understood that the evaluators' convictions and philosophies create a lens through which data are viewed, collected and interpreted. The starting point for the Knight study was critical discussion of LDPs in the press, and testing to see if LDP graduates shared these views. Could this critical starting point indicate the researchers' philosophical opposition to LDPs? Equally, does the position of Coffey and Gibbs, and of Prosser et al, lead them to an inevitable set of findings, conditioned by their own starting points?

Despite this note of caution, it is clear that large-scale evaluations have value in that they provide a useful back-drop for the local evaluations which, this paper asserts, are more directly useful for individual institutions. Multi-institutional studies provide tentative benchmark data, and make important links between LDP participants' experiences, and theoretical constructs. They draw the broad brush strokes of key issues, such as the nature of professional learning in the HE context (Knight, 2006); the question of how to balance generic teaching and learning with discipline-specific development (Prosser et al, 2006); and the important focus on student learning (Gibbs & Coffey, 2000). Since large studies lack institutional contextualisation, they need to be viewed through the filter of local profiles. Hence the Prosser et al finding (2006, p. 36) that participants in post-1992 institutions, which are more likely to be teaching-focused, perceived the programmes more positively than participants in the more research-focused pre-1992 institutions alerts developers to the need to design their evaluation within their local institutional context.

Learning from published, smaller scale evaluations

Institution-specific evaluations can offer some of the benefits of the larger scale studies if they are theory-informed and structured in approach, but with the major advantage of institutional relevance.

The first example of such an evaluation is of a programme which aimed to produce changes in participants' conceptions of teaching, towards approaches more conducive to student learning (Ho, 1998, p. 7). The evaluation methodology was designed to cover the effects of the programme on participants' conceptions, their teaching practices, and their students' approaches to study. This involved a longitudinal study, with semi-structured interviews at different stages. Students responded to the Course Experience Questionnaire (Ramsden, 1991) and the Approaches to Studying Inventory (Entwistle et al, 1979). Unusually, Ho established a control group, with which the programme participants compared positively. Most institutions do not have this option, since LDPs are a probationary requirement for new lecturers in most UK universities (Bamber, 2002). What can be learned from Ho, however, is the value of multi-stage evaluation, in order to chart participants' development over a period of time.

Another institution which based its LDP on a conceptual change / student-focussed approach is the University of Bangor. Lawson, Fazey and Clancy's (2007) evaluation administered two questionnaires, the Discipline Focus Epistemological Beliefs Questionnaire (Hofer, 2000) and the ATI, at programme beginning and end. Student data were also included, again using the Discipline Focus Epistemological Beliefs Questionnaire, to assess the relationship between teachers' beliefs and approach and students' approaches to study. A strength of this evaluation was the commitment to year on year collection of data, so that even relatively small initiatives might show some statistically significant results over time.

The LDP at Oxford Brookes University was predicated on a reflective practitioner model, with other, associated theories of change, such as the need for behavioural and conceptual development. Rust (1998, p. 256) described two small-scale evaluation studies, using an in-house end of course questionnaire, and externally-led interviews and focus groups. Although the evaluation did not link lecturer development to evaluations of improvement in student learning, as Gibbs and Coffey, and Ho, had done, it elicited lecturers' responses to the programme, using questions and statements designed to test their perception of whether better student learning had ensued. The learning from this evaluation is that local evaluations can gain objectivity from external facilitation, and that mixed methods can triangulate data.

The course described by Stefani and Elton (2002) had a different approach and theory of change: that adults learn best if involved in their learning, and see it as relevant to their needs. The distance learning programme aimed to convince teachers that university teaching was a problematic and researchable activity, and that a reflective, problem-based learning approach could aid development (Stefani & Elton, 2002, p. 118). In consonance with their theory of change, the evaluation approach used case studies to reflect on the course experience, together with data from the institution's internal quality audit. The lesson from this example is to seek alignment between the course and the evaluation approach.

Alignment was also key in a Queensland University of Technology study. The theory of change behind their LDP emphasized structured reflection as a cognitive tool for enhancing teaching effectiveness, and structured reflection was the evaluation method, with a theoretical framework to explicate the meaning of 'reflection' (Diezmann & Watters, 2006, p. 5). The reflective evaluation process was divided into its various components and scope (quick reading, zooming in, zooming out). This involved gathering data from a range of stakeholders, through open-ended surveys and through assignment work and unsolicited commentary, and these data were used to zoom in on the course experience. Zooming out involved strategic reflection within the institutional context, leading to general conjectures. In using these reflective tools, the causal simplicity of linking lecturer development with effective teaching and student learning could be avoided.

These small-scale evaluations are examples of locally-designed studies which are predicated on the theory of change of the particular course, and, as such, are more likely to be relevant to the institution. The danger is that local evaluation can lack externality, and careful triangulation is required to increase confidence in findings and cross-validate data (Jacobs, 2000, p. 271). The advantage is that these contextualised data can be understood and acted upon within the local situation, and evaluators can take a longer term, iterative approach. Another strength which the above studies share with the large-scale evaluations is the use of theoretical underpinning, related to a theory of change.

The answer to the question of what can be learned from small-scale, local studies, therefore, is that evaluation can be aligned with its context, and can provide insight into the longer term local impact of a programme. This can provide rich data, framed within a theoretical understanding which is shared with and understood by everyone involved, avoiding causal simplicity. The ensuing evaluation data inform further development of the course over time, and can lead to improved relationships between participants and the course team.

The parallels between published evaluations and the experience of a particular case institution

The following example of LDP evaluation in the author's own university will illustrate parallels with other published evaluations. The additional ingredient which may make this case study a helpful illustration for others is a structured evaluation framework which has been vital in framing the institution's longer term evaluation strategy. The LDP involved is the Postgraduate Certificate in Academic Practice (PG CAP), a 2-year, part-time, Masters level qualification, which is a probationary requirement for all new, inexperienced lecturing staff in a small, research-intensive university.

The starting point for the evaluation was a framework which requires evaluators to identify in advance the Reasons and Purposes of the evaluation, its Uses, Foci, Data and Evidence, Audience, Timing and Agency: RUFDATA (Reynolds & Saunders, 1987). To emphasise the importance of Context and Theories of Change, these factors were added to the framework by the PG CAP evaluators, so the tool became RUFDATACT. This tool was then used iteratively to review and update the evaluation approach over time. For example, the Teaching Methods Inventory (Coffey & Gibbs, 2002) was used and then dropped when it was found not to produce reliable data.

Focusing on the Context helped decide whether the evaluation should concentrate on a quantitative or qualitative approach. In a research-led university specialising largely in science, engineering and technology subjects, the technical-scientific orientation meant that quantitative evaluation data would be well-regarded. However, while quantitative methods seemed appropriate in this institutional context, and they can help to answer questions which hypothesise relationships between variables, the sample size (15-20 participants per year) would not give immediately significant results – if, indeed, significant results were obtainable. A longer term approach was necessary, in combination with qualitative methods which might better answer questions of meaning, interpretation and socially constructed realities (Newman et al, 2003, p. 170). Many social science researchers (e.g. Jacobs, 2000) advocate mixed methods, especially for complex environments like universities, with their pluralistic, contested cultures (Becher & Kogan, 1980, p. 55). The PG CAP evaluation adopted a mixed methods approach, with a combination of quantitative instruments (longitudinal questionnaires) alongside qualitative methods (interviews, open questions and reflective writing). However, the ontological and epistemological assumptions which underpin quantitative and qualitative methods are at variance (Bailey & Littlechild, 2001, p. 359), and so mixed methods evaluators need to clarify their assumptions. This is, in fact, helpful, rather than problematic, as a reminder to articulate both the theories of change of the course, and of the evaluation.

It might be assumed that the theories of change behind LDPs are obvious, since most (although not all) UK courses were initiated following the UK Dearing Report recommendation that lecturing staff should be trained, to promote better student learning (NCIHE, 1997, Para 70). In spite of this common policy directive, the goals and approaches of courses vary and are often multifaceted (Gibbs & Coffey, 2000, p. 32), and institutional programmes have their own flavour (Bamber et al, 2006). Theories of change are, therefore, unlikely to be straightforward for ambitious, context-rich programmes like LDPs. There were four theories of change underpinning the PG CAP programme, and each one was tested by different means.

The first theory of change (found also in most of the previously mentioned evaluations) stems from the extensive literature which links lecturers' conceptual frameworks to student outcomes (e.g. Kember & Kwan, 2000; Trigwell & Prosser, 1996; Trigwell, et al, 1999). The theory is that student-focused teachers can lead to better student learning. The quantitative tool used to test this change was the Approaches to Teaching Inventory (Trigwell and Prosser, 1999), administered on entry to and exit from the course, and, qualitatively, participants' reflective writing and interviews. Given the difficulty of distinguishing quantitative data trends with a small sample, the next stage of evaluation will plot this institution's ATI data against results obtained in other institutions, to check for comparisons. However, any comparison has to consider the influence of institutional context. The qualitative data has, so far, given more interesting results, as individuals explain how their attitude and approach to student learning have developed.

The importance of developing lecturers as reflective practitioners (e.g. Schön, 1983) is another common theory of change behind LDPs (Rust, 2000, p. 255), and the second PG CAP theory of change is that the reflective practitioner model of professional learning can help teachers to critically work through what this might mean for them. Quantitatively, the End of Course Questionnaire (adapted from Rust, 2000) aims to chart in what ways PG CAP participants feel they have changed as a result of the course, self-scoring on a Likert scale against a number of statements. This is an opportunity to test several levels of the theories of change, such as how helpful reflection is perceived to have been. Qualitatively, the Reflective Writing which is done throughout the course is specifically aimed at helping participants think critically about the relationship between theory and practical applications. However, using such data for evaluation purposes raises ethical and validity concerns if the participants are writing, as in this case, for assessment purposes. Participants are asked post hoc if relevant excerpts can be used for evaluation. Even so, data must be treated with care, as they may be writing what they think course tutors wish to hear. Triangulation with other sources is required to test whether these narrative data are valid.

Great value is attributed by new lecturers to the acquisition of practical skills which make them feel more confident and in control (Rust, 2000, p. 257). In the PG CAP evaluation, the third theory of change relates to the development of practical skills and confidence. The End of Course Questionnaire again tests participants' perception of their skills development, eliciting responses to statements like 'I am more skilled as a teacher of my discipline', and 'I am able to plan better courses and lessons as a result of this course'. Qualitatively, discussions and open written responses at the end of the course ask participants to comment on specific ways in which their teaching has developed, and to what they attribute that development.

In the learning literature, motivation is a significant factor in learning (e.g. Biggs, 1999; Boulton-Lewis, 1996; McAlpine & Weston, 2000; Vallerand, et al, 1992), and the fourth theory of change postulates that the programme will have different effects on individuals with different sources and levels of motivation, and perhaps from different subject groups. The initial PG CAP evaluation hypothesized that motivation levels of probationary ‘conscripts’ could be a key learning factor, and so a Motivation Questionnaire was designed. Like the ATI, this questionnaire is completed on programme entry and exit, to explore how attitudes to the programme have developed, and to chart disciplinary differences. Participants indicate their reasons for doing the course, against ten statements analysed into Extrinsic, Intrinsic or Strategic reasons. An extrinsic reason, for example, is that a participant attended the course “Because my head of department told me to” or “To pass probation”. An intrinsic motivation would be suggested by “Because I want to become a better lecturer”. A strategic motivator would be “To help me move on in my career.” The questionnaire also elicits the level of motivation for participating in the programme (high, medium, low), and changes to motivation between programme start and end. Results can be cross-tabulated with subject discipline and ATI results, and in the next stage of the evaluation it is hoped that the cumulative sample size will be sufficient to plot this relationship. Qualitatively, participants’ reflective writing also provides comments and stories which illustrate different experiences of the LDP, and confirm the need for further sub-group analysis.

Discussion

It appears that the large-scale approach to evaluation offers high level, high volume data coverage and, potentially, an objective external perspective, but may produce outcomes of limited relevance for specific institutions, with their own history, identity and needs. Smaller scale evaluations, on the other hand, focus on particular institutions, allowing for closer attention to detail and appreciation of the local environment. While the former view may serve wider, political purposes, the paper suggests that ‘zooming in’ on the local context (Diezmann & Watters, 2006, p. 5) has far greater potential for meaningful evaluation. The paper advocates a purposeful approach which acknowledges local strengths and needs, and also addresses the nature of lecturer development programmes (LDPs): they are not simply academic courses – they are change initiatives. The theory of change underlying the programme needs to be articulated. If the theory of change is tacit, as it often is (Trowler & Bamber, 2005, p. 81), then a useful opening discussion for the evaluation will establish what the LDP is specifically trying to change and how. Just as importantly, *why* the change works should be asked, to avoid the temptation of meeting the current political agenda, ‘knowing’ more but ‘understanding’ less (Weiss, 1995, p. 88). The course team can then tackle any evaluation findings which imply that the change initiative is not effective, for whatever reason, and the theory of change can be adapted. Their next iteration of the change initiative will be to adjust expectations, and reformulate the change theory: this would be difficult with a broad brush approach.

What has emerged, both from study of other published evaluations and from eight years’ experience of taking a structured approach to evaluating a specific programme in the author’s university, is that evaluation is a complex, iterative process. Evaluators can fall into many pits: for example, the ethical pit of researching the experiences of course participants who are simultaneously colleagues, probationers, and students who are presenting work for assessment; the epistemological pit of opting for research approaches which are not easily congruent with the theory of change of the LDP, or with the nature of small scale social science enquiry; the moral and political pit of needing data to justify the existence and continuation of provision; the workload pit of undertaking extensive evaluation research with no resource. Evaluation is clearly a serious activity. In spite of these caveats, this paper argues that the self-knowledge obtained by course teams who engage with local evaluation processes is significant, and that lecturer development programmes benefit from the effort. The paper aims to support developers in that effort, not by requiring them to carry out over-complex and time-intensive evaluation, but by helping them to frame their evaluations efficiently within a purposeful framework.

A number of lessons are apparent from the evaluations described above, which may be illustrative for other evaluators. First, evaluators need to gain credibility and relevance by working within the local culture. For example, quantitative research is an important part of the ‘neighbourhood culture’ in the author’s university, and so quantitative methods were tried there, despite the small sample size. Second, they should articulate the theories of change for the programme, and design the evaluation accordingly. This can help to focus the evaluation, but will also encourage the course team to exchange ideas about their theories of change – a useful dialogue. Then, evaluation plans can be formulated within a structured framework, such as RUFDATACT, underpinned by relevant evaluation literature.

Where larger studies can provide a backdrop or comparators for an evaluation, evaluators should use them, but focus on the characteristics of the institution. Once data have been gathered, simplistic interpretations of complex situations should be avoided: learn to live with uncertainty and ambiguity. This will involve a longer term approach; although the case study institution has been collecting data for eight years, programme outcomes may only become apparent over time. Data collection and interpretation is iterative, and the data become richer on each iteration, so evaluators need to be open to new findings. Finally, course teams should not underestimate the power of evaluation: even if data are not overwhelmingly significant, both learning and credibility can be gained.

Conclusions

This paper has sought to convince educational developers of the value of committing to their own version of theory-informed evaluation, and of publishing it to build a body of contextually-relevant knowledge of how LDPs work in their specific contexts. For the context in question, evaluation in a local setting, informed by theory and other empirical studies, has made several contributions. There has been impact on university management, who respect the research-informed and research-oriented evaluation process; impact on the programme team, who use the data to update the course, and are considered a 'listening team'; and impact on course participants and the programme itself, which is constantly being reviewed and questioned in the light of data obtained.

These gains have been made with reference to the theories of change for the course. The focus of future research stems from emerging findings, that positive things happen during the programme for most – but not all – participants, and that subject discipline and departmental support appear to be factors. An important future research question, then, is not "Does the programme work?", but "What in the programme works for whom, and why?" (Pawson & Tilley, 1997, p. 109, 113). Formulating insightful answers to these questions requires cumulation (Pawson & Tilley, 1997, p. 116), and so the evaluative research will continue. Given small numbers, continuation is likely to involve comparison with other institutions: this may illuminate which aspects of contextual specificity are influential. Another important evaluation strand relates to longitudinal impact, and how the longer term effect of a course like PG CAP can be ascertained, if participants meet departmental opposition to change and only become influential perhaps five years after completing the programme. In the meantime, no claims are made about certainty of current findings, but institutional reference points have been understood, insights have been obtained and the research will continue.

Acknowledgements: I am grateful to Professors John Cowan, Paul Trowler, Peter Knight and Drs John Bamber and Karen Smith for their insightful comments on previous drafts of this paper.

References

Bailey, D., & Littlechild, R. (2001). Devising the evaluation strategy for a mental health training programme. *Evaluation, 7* (3), 351-368.

Bamber, V., Walsh, L., Juwah, C., & Ross, D. (2006). New lecturer development programmes: A case study of Scottish higher education institutions. *Teacher Development, 10*(2) , 207 – 231.

Bamber, V. (2002). To what extent has the Dearing Policy Recommendation on Training New Lecturers met acceptance? Where Dearing went that Robbins didn't dare. *Teacher Development*, *6*(3), 433-457.

Becher, T., & Kogan, M. (1980). *Process and structure in higher education.* London: Heinemann.

Biggs, J.B. (1999). *Teaching for quality learning at university*. Buckingham: SRHE and Open University Press.

Blackler, F. (1995). Knowledge, knowledge work and organisations: An overview and interpretation. *Organisation Studies,* 16, 6, pp 1021-1046

Boulton-Lewis, G.M., Wilss, L., & Mutch, S. (1996). Teachers as adult learners: Their knowledge of their own learning and implications for teaching. *Higher Education, 32*(1), 89-106.

Coffey, M. & Gibbs, G. (2002) Measuring Teachers' Repertoire of Teaching Methods. *Assessment & Evaluation in Higher Education.* Vol 27, No 4, pp 383-390

Coffey, M., & Gibbs, G. (2000). The strategic goals of training of university teachers. *Paper from 8th International Improving Student Learning Symposium*, UMIST, Manchester

Diezmann, C.M., & Watters, J.J. (2006). Structuring reflection as a tool in qualitative evaluation. *Paper presented at HECU Conference, Lancaster*. http://www.lancs.ac.uk/fss.events/hecu3/documents/diezmann_watters.doc (Last accessed 1 September 2006).

Entwistle, N., Hanley, M. and Hounsell, D. (1979) Identifying distinctive approaches to studying. *Higher Education*, Vol 8, No 4, pp 365-380.

Gibbs, G. and Coffey, M. (2004) The Impact Of Training Of University Teachers on their Teaching Skills, their approach to teaching and the approach to learning of their students. *Active Learning in Higher Education*. 5, 87-100.

Gibbs, G. and Coffey, M. (2000) Traiing to teach in higher education: a research agenda. *Teacher Development.* Vol. 4, No. 2, p 31-44.

Gosling, D. (2008). *Educational development in the United Kingdom. Report for the Heads of Educational Development Group.* HEDG: http://www.hedg.ac.uk/documents/HEDG_Report_final.pdf (Last accessed 3 March 2008).

Ho, A.S.P. (1998). A conceptual change approach to staff development: A model for programme design. *Paper presented at 2nd ICED Conference, Austin*. http://www.uni-bielefeld.de/IZHD/ICED/Ho-Program.html (Last accessed 12 September 2006) [Also published in Ho, A.S.P. (2000) A conceptual change approach to staff development: A model for programme design. *The International Journal for Academic Development, 5* (1), 30-41]

Hofer, B. K. (2000) Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology*. 25, pp 378 – 405.

Jacobs, C. (2000). The evaluation of educational innovation. *Evaluation*, 6(3), 261-280.

Kember, D. and Kwan, K.P. (2000) Lecturers' approaches to their teaching and their relationship to conceptions of good teaching. *Instructional Science*, 28: 469-490.

Knight, P. (2006). The effects of postgraduate certificates: A report to the project sponsors and partners. Milton Keynes: The Open University.

Lawson, R.J., Fazey, J.A., & Clancy, D.M. (2007). The impact of a teaching in Higher Education scheme on new lecturers' personal epistemologies and approaches to teaching. In C. Rust (Ed.), *Improving student learning*. Oxford: Oxford Centre for Staff Development.

McAlpine, L. (2000). Training university teachers: Theory and evidence. *ICED Symposium* http://www.uni-bielefeld.de/IZHD/ICED/mcalpine.html (Last accessed 12 September 2006).

McAlpine, L., & Weston, C. (2000). Reflection: Issues related to improving professors' teaching and students' learning. *Instructional Science, 28*(5-6), 363-385.

Marsh, H. W. (1982) SEEQ: A Reliable, Valid, and Useful Instrument for Collecting Students' Evaluations of University Teaching. *British Journal of Educational Psychology*, Vol 52, Part 1, pp 77-95.

Martin, E. (1999). *Changing academic work: Developing the learning university*. Buckingham: SRHE and Open University Press.

NCIHE (National Committee of Inquiry into Higher Education) (1997). *Higher education in the learning society*. Committee chaired by Sir Ron Dearing. Available at http://www.leeds.ac.uk/educol/ncihe (last accessed 31 July 2007).

Newman, I., Ridenour, C.S., Newman, C., & DeMarco, G.M.P. (2003). A typology of research purposes and its relationship to mixed methods. In A. Tashakkori & C. Teddlie (Eds), *Handbook of mixed methods in social and behavioral research*. (pp. 167-188) Thousand Oaks: Sage.

Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London: Sage Publications.

Prosser, M., Rickinson, M., Bence, V., Hanbury, A., & Kulej, M. (2006). *Report: Formative evaluation of accredited programmes*. York: The Higher Education Academy.

Ramsden, P. (1992). *Learning to teach in higher education*. London: Routledge.

Ramsden, P. (1991) A performance indicator of teaching quality in higher education: the Course Experience Questionnaire, *Studies in Higher Education*, 16, pp 129-50

Reynolds, J., & Saunders, M. (1987). Teacher responses to curriculum policy: Beyond the 'delivery' metaphor. In J. Calderhead, (Ed), *Exploring teachers' thinking*. (pp. 195-213). London: Cassell Education.

Rust, C. (2000). Do initial training courses have an impact on university teaching? The evidence from two evaluative studies of one course. *Innovations in Education and Training International, 37*(3), 254-261.

Rust, C. (1998). The impact of educational development workshops on teachers' practice. *The International Journal for Academic Development, 3*(1), 72-80.

Schön, D. (1983). *The Reflective Practitioner. How professionals think in action*, London: Temple Smith.

Stefani, L., & Elton, L. (2002). Continuing professional development of academic teachers through self-initiated learning. *Assessment and Evaluation in Higher Education*, 27(2), 117-129.

Trigwell, K., Prosser, M., & Waterhouse, F. (1999). Relations between teachers' approaches to teaching and students' approaches to learning. *Higher Education, 37*, 57-70.

Trigwell, K., & Prosser, M. (1996). Changing approaches to teaching: A relational perspective. *Studies in Higher Education, 21*(3), 275-284.

Trowler, P., & Bamber, V. (2005). Compulsory higher education teacher education: Joined-up policies; institutional architectures; enhancement cultures. *International Journal for Academic Development, 10*(2), 79-93.

Vallerand, R.J., Pelletier, L.G., Blais, M.R., Briere, N.M., Senécall, C., & Vallieres, E.F. (1992). TheAacademic Motivation Scale: A measure of intrinsic, extrinsic and amotivation in education. *Educational and Psychological Measurement, 52*, No.4, 1003-1019.

Weiss, C.H. (1995). Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families. In J. Connell, et al. *New approaches to evaluating community initiatives: Concepts, methods, and contexts*. (Vol. 1, pp. 1-16). Washington, DC: Aspen Institute.