

Learning about Second Language Assessment: Insights from a Postgraduate Student On-line Subject Forum

KIERAN O'LOUGHLIN

University of Melbourne

This article aims to examine student learning in a postgraduate subject on second language assessment. This is a notoriously difficult domain of knowledge for students in second language teacher education programs because of the high level of abstraction around its key theoretical concepts, validity, reliability, and practicality, and how they need to be balanced against each other in designing and using assessment instruments. Asynchronous on-line forums, used as adjunct learning environments to the classroom, offer rich insights into how individual students grapple with complex course content of this kind. Using print-outs of selected weekly forum discussions as data, it is shown that students' capacity and willingness to grasp the new knowledge and ideas presented to them in class are strongly influenced by their prior experience with assessment as language learners and teachers as well as the quality of the input they receive in class sessions. The implications of the findings are examined for in-service teacher education programs more broadly.

INTRODUCTION

Postgraduate programs in second language teacher education often include a core or elective unit on assessment. This is an increasingly important domain of language teachers' expertise as the professional demands on them to accurately assess their

Address for correspondence: Kieran O'Loughlin, Doug McDonell Building, Faculty of Education, The University of Melbourne, Victoria 3010, Australia; Email: kjo@unimelb.edu.au

University of Sydney Papers in TESOL, 1, 71-85.

©2006 ISSN 1834-3198 (print) ISSN 1834-4712 (online)

students increases and as the theory and practice of assessment continues to mature. However the literature on this specific area of second language teacher education is surprisingly limited.

A survey conducted by Bailey and Brown (1995) of 84 faculty staff who taught a course in language testing revealed that test critiquing was covered the most overall, followed by item writing, test score interpretation, test revision, test scoring, test administration and, lastly, test taking. The questionnaire they used was oriented towards the more traditional testing paradigm with the majority of questions focusing on standardised tests and statistical procedures related to item analysis, test consistency and test validity.

It is only very recently that there has been a change of emphasis from language "testing" to the more inclusive term "assessment" in teacher education courses to reflect a shift in perspective across the field. This means that formal testing is now recognised as only one kind of assessment and that the training needs of testing experts and teachers in this area are not the same in so far as teachers use a variety of formal and informal methods of assessment in the classroom for placement, diagnosis, formative and summative purposes. This shift has also been accompanied by the realisation that assessment is not simply a technical, psychometric area of knowledge. Instead, like all aspects of teaching, it is now better understood as a social practice with its own complex political and ethical considerations.

Teachers, it seems, are mainly concerned with the relationship between assessment and learning. Based on a survey of studies in a range of educational contexts, Brindley (2001: 127) suggests that "teachers see assessment as an activity which is integrated into the curriculum with the aim of improving learning, rather than a 'one-off' summative event." Brindley (2001: 129) argues that professional development programs aimed at classroom practitioners should 1) focus on curriculum-related assessment; 2) exploit teachers' existing knowledge; and 3) be adaptable to meet a wide variety of teacher needs. The essential components of such a program would include the following units: the social context of assessment, defining and describing proficiency, constructing and evaluating language tests, assessment in the language curriculum and putting assessment into practice. Brindley (2001) suggests that these professional development programs should encourage teachers to see how their assessment practices fit into the broader educational context, capitalise on their existing practices, recognise and deal with the reality and constraints of their assessment practices, encourage a research orientation to professional development and help teachers plan for change.

Shohamy (2005: 107) also argues that professional development in assessment is not a question merely of demonstrating the technical 'tricks of the trade'. She argues for "... the

need to expand the role of teacher education programs in which teachers are exposed not only to the procedures and methods of testing and assessment but also to aspects related to the consequences of tests." From a critical testing perspective Shohamy (2005: 108) suggests that " [t]eachers need to become more aware, more socially responsible and socially reflexive about the uses of tests by pointing out such misuses to the public at large and by resisting the 'one size fits all' approach."

While Brindley (2001) and Shohamy (2005) indicate the appropriate directions for teacher education programs in assessment to take, there appears to be a lack of recent empirical research into the content of such programs and, more particularly, how teachers engage with and integrate the new knowledge presented to them. A notable exception is a study by Kleinsasser (2005) which describes the evolution of an innovative postgraduate assessment and testing course in which the main task for participants during the semester was to develop assessment materials informed by the theory covered in the classes. The methodological approach adopted by Kleinsasser (2005) was narrative enquiry which has recently been adopted by a number of researchers in second language education to explore teachers' experiences (see, for example, Bailey & Nunan, 1996; Freeman & Richards, 1996; Freeman & Johnson, 1998; Contreras, 2000; Johnson & Golombek, 2002; Golombek & Johnson, 2004). However, as Kleinsasser (2005: 79) notes, there has been a dearth of teacher educator narrative inquiry. In Kleinsasser's (2005) own study the unfolding of his assessment and testing course is presented mainly through the author-as-lecturer's own eyes (although the students' final evaluations of the subject are also included). The current study also attempts to map the learning from the lecturer's perspective. However, in this case, the narrative is grounded in students' on-line forum contributions. This approach allows for a close examination of how the students grapple with new knowledge and ideas presented to them in class over the course of the semester.

CONTEXT OF THE STUDY

The subject which provides the context for this study is called "Assessment in the Language Classroom". It is offered as a postgraduate elective in TESOL and Modern Languages programs in the Faculty of Education at the University of Melbourne. The subject consists of 36 hours of face-to-face instruction (three hours per week over a 12 week semester). The subject aims to enable students to develop (a) a sound understanding of key concepts in second language assessment; (b) their ability to critically evaluate existing assessment instruments, and (c) their capacity to design or adapt assessment instruments for their particular teaching contexts. The following topics are included: introduction to second language assessment, theoretical and practical aspects of designing assessment tools, assessing speaking, listening, reading and writing,

evaluating assessment tools, self and peer assessment, integrating assessment and learning, policy and social issues in language assessment, current trends and future directions. In all of the class and online forum sessions participants were encouraged to apply the new concepts and ideas presented to them in class to examples of language assessment practice including their own.

Study Participants

Twelve students took part in the study: five local and seven international students. The international students came from China (4), Fiji (1), Vietnam (1), and Thailand (1). There were eleven female and one male students enrolled in the subject. Eleven of them were qualified in TESOL, three were also accredited teachers of either Chinese or Italian, and there was one qualified speech pathologist. All of the students had at least one year of teaching experience.

The Online Forum

An asynchronous online forum was offered as an adjunct to the classroom sessions. In this kind of forum students participate by posting written contributions in their own time. In the subject, "Assessment in the Language Classroom", students were required to make regular written contributions to the forum as a hurdle requirement i.e., their active participation (at least one substantial posting per week) was obligatory although it was not assessed as part of the formal assignment work. Each student had to facilitate one week's discussion and be a respondent in the other weeks. The forum was led by the lecturer (the author of this paper) during the first week and then progressively by each member of the class in subsequent weeks. The facilitator posed an initial question or issue arising from the previous class and then guided the discussion and summarised it at the end of the week. The other contributors were required to make weekly contributions by responding to the facilitator's question or problem and/or engaging with the other comments which followed. They were also able to raise other relevant issues as the week progressed. The forum's asynchronous format enabled students to participate at times that suited them between the weekly class sessions.

After facilitating the first week's discussion I monitored student contributions in weeks 2 to 12. However, I progressively withdrew from active participation intervening only where there was a clear breakdown in the group's understanding. At the end of the subject I checked again to ensure that all students had made regular contributions throughout the semester.

Online forums of the type used for this subject offer several potential benefits. They (a) provide students with the opportunity to actively reflect on the preceding class

sessions in an interactive environment, (b) help create and support collaborative learning, and (c) enable the lecturer to evaluate how students are coping with the course content and how effectively they have been taught (Andriesson, Baker & Suthers, 2003; Naidu, 2003; Robbins & Webster, 2003; Love & Iles, 2006).

APPROACH

The forum contributions were printed out each week and monitored for the quantity and quality of students' contributions overall. In order to examine how individual students engaged with the content of the subject the contributions of two students were also tracked over the 12-week duration of the course - one international ('Wei Yan') and one local student ('Roula'). These students were selected because they were both regular and substantial contributors to the forum. It was therefore possible to monitor their learning in more detail than most of the other students whose contributions were more brief and/or sporadic. The findings are presented as parallel narratives which highlight the similarities and differences in the students' learning in the subject.

FINDINGS

General Findings

An initial examination of the contributions during the twelve weeks the forum ran suggested that it was fulfilling the functions outlined above, i.e. it enabled students to actively reflect on the class sessions in an interactive environment, it helped create and sustain collaborative learning, and it enabled the lecturer to evaluate student learning as well as the quality of his own teaching.

Most students made fairly regular contributions to the forum discussions although only a few participated actively in all twelve discussions. Their contributions tended to be either introspective and/or retrospective in character, making useful and sometimes highly insightful links between the course content and their own experiences as both language learners and teachers, i.e., as assessed students and assessing teachers. However, their contributions also highlighted the interactive value of the forums: students provided supportive but also, at times, critical feedback to each other as they attempted to merge their previous and new understandings about assessment. This was not always a simple process where new ideas presented in class were automatically digested and unreservedly accepted. Rather, the forum enabled students to establish and debate their evolving views on assessment and receive valuable feedback from their peers in so doing.

In the end-of-course evaluation questionnaire most participants indicated that they found this activity an extremely valuable adjunct to the face-to-face classes. However, some students reported having trouble gaining access initially to the forum because of technical problems. Some of them also had no experience of using this kind of communication tool and took a few weeks to participate effectively. It was also notable that international students, for whom English was an additional language, contributed probably more than local students to the forum, while the reverse was true in class. This kind of communication may suit second language learners more than 'real time' oral classroom discussion as it allows for responses which are more planned. Finally, while there was carry over of content from the class to the forum the reverse rarely happened - the forum discussions remained a series of communicative events which did not impact on the class sessions, at least directly.

Case Studies

Wei Yan was a Chinese international student who had taught for 9 years in the English department of a university in Beijing. She was particularly interested in assessing speaking. She had a wide range of experience in large-scale speaking tests as both an interviewer and assessor. She had been an assessor for speaking tests conducted in her own university as well as the oral component of the National College Entrance Exam in China. She had also recently taken the IELTS (International English Language Testing System) test to gain entry to her current postgraduate course.

Roula was a local Australian student of Greek background with 5 years teaching experience as a generalist primary teacher prior to taking up a position as an ESL teacher for newly arrived immigrant students nine months before she took this subject.

The contributions made by Wei Yan and Roula over selected weeks of the 12-week semester form the basis of the analysis which follows:

Week 1: Introductions

Wei Yan introduced herself in week 1 by describing the pressures she faced as an examiner of the spoken English component of the standardized, highly competitive National College Entrance Exam in China:

I sometimes had more than 80 students come in one day for the oral interview, which was just awful. But I made it.

She was also quick to admit her limitations at this early point in the subject:

I should say I've got a bit [of] experience with language assessment, but quite superficially. I do need to do some solid theoretical work in this class.

Roula's experience in assessment was much more of the classroom-based kind:

I use a variety of assessment techniques including ...a lot of observation and oral (informal) assessment.

Like Wei Yan, Roula had clear motivation for doing the subject, in her case to gain more knowledge of assessment techniques:

I'm not very confident in the area of assessment and would like to learn more about appropriate methods of assessing, which is why I chose this subject.

Week 3: Assessment - Practical Concerns

In her contribution Wei Yan for week 3 outlined how there had been a huge upsurge in the number of students taking the speaking component of the National College Entrance Exam in mainland China in 2000. She described how candidates, accompanied by their parents, formed long queues to be interviewed. Some of them waited nearly the whole day to be tested. Many of them became extremely frustrated and lost their temper. The interviewers, in groups of two, who had to conduct the test with 50 to 70 students individually per day in previous years, now had to examine 150 students on that day, working from 7am to 6pm without a break. The interviews had to be shortened from 7-9 minutes to 3-5 minutes to accommodate this number of students. Not surprisingly, many of the interviewers became exhausted and gradually lost their patience. Wei Yan clearly understood the negative implications of shortening the test in the interests of practicality, particularly highlighting the issue of intra-rater reliability:

... the administration of the speaking test in 2000 was problematic due to the organisers' wrong expectation of the number of candidates, which thus made the test a lot less reliable and valid. Dear classmates, if you worked from 7am till 6pm and interviewed 150 students a day, would it be possible to be consistent in your marking?

Roula's assessment work was much less formal and stressful, carried out within the confines of her own classroom. For her, practicality was a less pressing concern: She believed that, in order to get a clear indication of a student's communicative competence, a number and variety of assessment instruments (both formal and informal), need to be administered. She favoured the ongoing use of assessment through the entire course even if some of the assessment tools were not very practical and not all students could be assessed at the same time. She argued that:

Variety is the spice of life... and there should be many and varied assessment tasks for the many and varied individual learners.

She did acknowledge, however, that this is more possible with small numbers of students.

The contrast here between Wei Yan and Roula's perspectives here underscores the two very different cultures of assessment and the situations in which they worked as assessors. Clearly, the conditions under which Roula operated in Australia were more conducive to classroom-based assessment which was ongoing and varied. This kind of assessment enjoys greater legitimacy in the Australian educational context than its Chinese counterpart where externally-designed standardised testing is used more exclusively for many assessment purposes including university selection. Of course, the difference between the students' viewpoints also reflected the different educational sectors in which they worked i.e. university versus primary school.

Week 4: Assessing Speaking

In the week 4 forum discussion Wei Yan was still focused on issues related to formal, large-scale standardised testing. She argued that:

Our best efforts should be invested to reduce the subjectivity and increase the reliability (of our marking).

For high stakes speaking tests she stressed the need for regular rater training, for more than one rater for each performance, and recordings of the performances so that both candidate and rater performance could be checked later on.

Roula's experience with assessing speaking was very different from Wei Yan's. In her role as classroom teacher she claimed:

I have never given the students 'speaking tests'... my assessment is purely observation and very subjective.

Roula discussed the students with their mainstream teachers on a regular basis "so there are in essence two teachers assessing the students, subjectively." Unlike Wei Yan she was resistant to the notion that her assessment should be as 'objective' as possible.

I really don't know how (assessing) speaking can be more objective and, in a primary classroom, I'm not sure that children need to be given more objective speaking tests.

Again, there are two very different views about assessment foregrounded here, the first aspiring to objectivity and the second rejecting that notion. Both are appropriate given the nature of the different assessment procedures and the contexts in which they take place: reliability becomes more crucial as the stakes of assessment become higher.

Week 5: Assessing Listening

In the week 5 forum on assessing listening Wei Yan rather surprisingly contributed the following viewpoint:

What has been discussed so far relates more closely to what I would call 'listening tests or exams', rather than assessment. Shall we discuss it in a broader sense under the term 'assessment'?

Her comment marked a turning point in her forum contributions, a re-orientation of her attention away from the narrower concept of "testing" to the more inclusive one of "assessment". She continued in the same vein:

For instance, shall we employ self-assessment during the semester to help the students to assess their own listening skills and monitor their progress?

As well, she suggests the use of portfolios to "keep track of the students' progress". Wei Yan's contribution here demonstrates her willingness to take on a new position in relation to a fundamental but (for her) new distinction between 'testing' and 'assessment'.

Roula's contribution, on the other hand, reflected her continuing opposition to "standardised testing":

A high stakes test such as the IELTS (listening test) would be...very stressful to second language learners.

In so doing she underscored the advantages of ongoing low-stakes classroom assessment:

...it may be a better idea to have a number of short listening tests spread over time rather than one big one which would undoubtedly increase the pressure placed on students. This may not be as practical, but I think the validity and reliability in this instance is more important.

Roula saw clearly the strengths of classroom-based assessment but also continued to oppose standardised testing. In this sense, she displayed a more rigid stance than Wei Yan who showed evidence here of being open to different kinds of assessment.

Week 8: Evaluating Assessment Instruments

By Week 8 both learners reached a crucial stage in their learning. Wei Yan had now gained the knowledge and confidence to be able to evaluate the assessment instruments handed down by educational authorities:

In the past I probably even won't give a thought to challenge the validity, reliability and practicality of some very important or high-stake tests... 'cause they are sort of given by authority. From last week's session and the previous ones I learnt to examine the tests in a more critical way and professional way and realised that no test could be perfect.

She had therefore fulfilled one of the subject's central aims:

I think the best point of this course is to train me to be critical.

Roula too appeared to be meeting this goal:

Just like most of you I too am learning to be more critical of language tests and really found the last class very enlightening.

Both students appeared to have benefited from the close evaluation of a formal ESL examination paper for final-year secondary students which we carried out in class that week. Roula claimed:

It's hard to believe that the sample Kieran gave us of a 'high-stakes' test can have so many problems in terms of validity and reliability. At first glance, the test seemed fair and I think this highlights how difficult it is to design a good assessment instrument.

However, Wei Yan showed more evidence of progressing in her thinking. She demonstrated in week 5 that she was open to re-thinking the definition of assessment and here she displayed the (for her new) readiness to critique standardised tests in a balanced way. Roula's comments were more predictable in so far as they were consistent with her opposition to standardised testing.

In retrospect, it would have been useful to have critiqued a more informal classroom-based procedure in this week as well to ensure that participants understood that *both* standardised tests and classroom-based assessments were open to critical scrutiny although the criteria for judging them may be different.

Week 9: Self and Peer Assessment

The week on self and peer assessment is often very confronting for students in the subject because it challenges their notions about teacher authority in the classroom. In the forum the two students were both positive about self assessment but Roula was less enthusiastic about peer assessment. Wei Yan reflected thoughtfully on the importance of learner training in both these types of assessment:

... students need some training on self-assessment and peer assessment... it is important to tell the students why we do so at the beginning of the training. Only when they realise that they themselves will be benefited will they be really motivated and involved. Otherwise, they would just do it for the teacher as part of the requirements.

She also understood the cultural challenges of adopting less teacher-centred assessment procedures:

...cultural background should also be taken into account. Students from Eastern cultures might be too shy to assess others. What's more the traditional role of the teachers and students will also influence the students.

Roula, too, was open to experimenting with self-assessment after the class session:

I found this weeks' session...most interesting and am for the first time going to try self-assessment with my students this week.

However, she was clearly influenced by another student's negative account of attempting to use peer assessment:

I found X's experience with peer assessment very interesting ... It was good of her to share a 'real-life' experience. Although I can see the benefits of self-assessment as an addition to other forms of assessment, I'm not as yet convinced of the virtues of peer assessment.

Week 10: Integrating Assessment and Learning

The week on integrating assessment and learning required students to consider assessment in relation to their own teaching as well as learner progress. By this point in the subject, Wei Yan clearly understood the multiple roles that assessment can fulfil:

It seems that every session we have had makes me think about assessment in a broader sense. This week's class made me realise that the results of assessment are not only 'indicators of student achievement' but can serve other purposes.

Roula articulated the same understanding in greater depth, although again there was less sense of her having progressed in her thinking given her bias towards classroom-based assessment:

Like other students I now have a clearer appreciation of the importance of integrating assessment and learning.... The notion of assessment as a way of evaluating teaching and learning isn't (I believe) often used. I feel that the focus on teaching and learning should be the first priority, with on-going formal and informal assessment as a means of improving teaching and learning.

Both students appeared to be very positively disposed to the concept of using information from student assessment for purposes other than gauging student achievement.

Week 11: Policy and Social Issues in Assessment

In the week on policy and social issues in assessment we spent some time in class discussing the ethical dilemma of whether or not teachers have the right to fail students, given the impact such a decision may have on their life chances. The debate carried over into the forum. Wei Yan contributed to the discussion as follows:

Do we teachers have the right to fail students? I would argue that students can't be failed only on their performance from a sole test or exam. The teachers have no right to fail the students from the results of one or two tests. However, the teachers would be able to fail a student who has shown poor performance during the whole semester. Otherwise it is not fair to others who work hard.

Roula, on the other hand, suggested that assessment should inform learning rather than being used to make definitive judgements about students:

I would never give students a pass or a fail, but instead utilise a variety of assessment tasks to provide students with detailed feedback on their progress, strengths and areas in need of improvement as well as useful strategies in order to improve language and learning in general.

With regard to 'high stakes' tests she argued that:

Test results can never really tell what a student is capable of doing. They merely show how well a set of particular questions were answered under pressure on one particular day.

Her comments were consistent with her unchanging opposition to testing, especially for making judgements about student's level of achievement. Roula then narrated how her

application to enter teachers' college was initially unsuccessful because of her relatively weak final-year secondary examination results. She was eventually admitted when other applicants decided not to take up their course offers. She felt that she was becoming a very successful teacher whereas her friend (with excellent examination results), who went into teaching for the wrong reasons ("because of the holidays"), had since left the profession. Her story provides more support for her lack of faith in testing in so far as it underscores the importance of personal factors such as professional suitability and motivation as opposed to examination results when tertiary institutions or employers make decisions about the suitability of an applicant for a course of study or a career.

DISCUSSION

The contributions of Wei Yan, Roula and the other students to the online forum indicated that the first two of the three main aims of the subject had been broadly achieved, i.e. participants developed a sound understanding of key concepts in second language assessment as well as their ability to critically evaluate existing assessment instruments. Evidence for the fulfilment of the third aim, for students to develop the capacity to design or adapt assessment instruments for their particular teaching contexts, was demonstrated through their final written assignment at the end of the semester.

The two case studies offer a clear contrast in terms of the individuals' cultural backgrounds and professional experience in assessment. Wei Yan's experience with standardised tests of speaking within the Chinese university context differed markedly from Roula's experience with classroom-based assessment in Australian primary classrooms. The two students also differed in their willingness to alter their preconceived views and extend their thinking. Wei Yan displayed a greater openness towards a broader, more complex view of assessment in general and to procedures with which she was less familiar (in her case classroom-based assessment), perhaps as a result of her negative experience with standardised testing. Roula, on the other hand, showed less progress in her thinking even though she started from a less traditional position supporting classroom-based assessment. She appeared to remain uncritically committed to that form of assessment while seeing no role for standardised tests at all, possibly as a result of her negative experience of testing as a student and more positive experience with classroom-based assessment as a teacher.

However, Roula's lack of progress may also have stemmed from a problem in the course design and delivery, namely, that more attention could have been given to building students' skills in critiquing classroom-based assessment procedures as well as standardised tests to ensure that all students arrived at a balanced view on the strengths and weaknesses of both kinds of instruments.

CONCLUSION

The online forum contributions in general and those of the two individual case study participants in particular, suggested that most students had developed a sound understanding of key concepts in second language assessment as well as their ability to critically evaluate existing assessment instruments. The individual case studies also showed that students' capacity and willingness to grasp the new knowledge and ideas presented to them in class were strongly influenced firstly, by their prior experience with assessment as language learners and teachers, and secondly, by the quality of the input they received in the class sessions.

These findings suggest that it is important for faculty staff to carefully consider their students' diverse cultural backgrounds, professional experiences and learning needs when planning and conducting their courses in second language teacher education programs. It is clear that participants' learning is influenced by their past experiences (both positive and negative) as language students and teachers. However, students are also responsible for recognising how the past has shaped their attitudes and for actively reflecting on that process as an integral part of their learning. The findings also emphasise the importance of ensuring that lecturers provide an even-handed perspective on key concepts and topics so that students can be exposed to, and hopefully develop for themselves, a critically balanced view of the domain of theory and practice about which they are being educated.

THE AUTHOR

Kieran O'Loughlin is Senior Lecturer in TESOL and Assistant Dean (International Programs) in the Faculty of Education at the University of Melbourne, Australia. He teaches a number of subjects on the postgraduate TESOL program including Second Language Assessment, English for Specific Purposes, Designing and Evaluating the English Language Teaching curriculum, and Research Methods in Language and Literacy Education. His main research interests include second language assessment (both standardised testing and classroom-based assessment), English for academic purposes, and social identities in education more broadly.

REFERENCES

- Andriesson, J., Baker, M. & Suthers, D. (Eds.) (2003). *Arguing to learn: confronting cognitions in computer-supported collaborative learning environments*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

- Bailey, K.M. & Brown, J.D. (1995). Language testing courses: what are they? In Cumming, A. & R. Berwick (Eds.), *Validation in language testing* (pp.236-256). Clevedon: Multilingual Matters.
- Bailey, K.M. & Nunan, D. (Eds.) (1996). *Voices from the language classroom*. Cambridge: Cambridge University Press.
- Brindley, G. (2001). Language assessment and professional development. In E. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. McNamara & K. O'Loughlin (Eds.), *Experimenting with uncertainty: language testing essays in honour of Alan Davies. Studies in Language Testing Volume 11* (pp.126-136). Cambridge: Cambridge University Press.
- Contreras, G.A. (2000). Self-storying, self-understanding: Toward a narrative approach to EFL teacher education. *TESOL Journal*, 9, 24-27.
- Freeman, D. & Johnson, K.E. (1998). Reconceptualising the knowledge-base of language teacher education. *TESOL Quarterly*, 32, 397-417.
- Freeman, D. & Richards, J.C. (1996). *Teacher learning in language teaching*. Cambridge: Cambridge University Press.
- Golombek, P.R. & Johnson, K.E. (2004). Narrative inquiry as a mediational space: Examining emotional and cognitive dissonance in second-language teachers' development. *Teachers and Teaching*, 10, 307-327.
- Johnson, K.E. & Golombek, P.R. (Eds.) (2002). *Teachers' narrative enquiry as professional development*. Cambridge: Cambridge University Press.
- Kleinsasser, R. C. (2005). Transforming a postgraduate level assessment course: A second language teacher educator's narrative. *Prospect*, 20, 77- 102.
- Love, K. & Iles, M. (2006). 'Welcome to the online discussion group': Towards a diagnostic framework for teachers. *Australian Journal of Language and Literacy*, 29, 210-225.
- Naidu, S. (Ed.) (2003). *Learning and teaching with technology: Principles and practices*. London: Kagan.
- Robbins, K. & Webster, F. (Eds.) (2003). *The virtual university*. Oxford: Oxford University Press.
- Shohamy, E. (2005) The power of tests over teachers: the power of teachers over tests. In D. J. Tedick (Ed.), *Second language teacher education* (pp.101-111). Mahwah, NJ and London: Lawrence Erlbaum.