

Transforming a postgraduate level assessment course: A second language teacher educator's narrative

ROBERT C KLEINSASSER – The University of Queensland

ABSTRACT

The article describes an experience changing a postgraduate Applied Linguistics and TESOL Studies assessment and testing course. A (second language) teacher educator's narrative relies on descriptions and reflections of the course, course outline, and course processes, along with comments made by students in their end of semester evaluations. The study is situated within the growing literature of (second language) teachers' and (second language) teacher educators' narrative inquiry. Background of the course is offered and discussion of two specific processes in professional development is presented. Description of course members negotiating assessment content and assessment materials development is then narrated. The implementation of a portfolio process is also tendered. In essence, the manuscript attempts to share three stories at once: one about assessment, one about pedagogy, and one about trajectories of professional participation through a (second language) teacher educator's narrative.

Introduction

In this paper, I construct a narrative of how my students and I began transforming a postgraduate, second language assessment course. Drawing on my descriptions of, and reflections on, the course, course outline and course processes, along with comments made by students in their end-of-semester evaluations, I also report on the changing manner in which we, ourselves, learned and developed.

First, I situate the study within the growing literature of narrative inquiry, reviewing this particular genre in the areas of (second language) teachers and (second language) teacher educators. Second, I offer a background to the course, giving an idea of previous course content and pedagogy. Third, I discuss two specific processes in professional development. Fourth, I describe how we, as a class, began negotiating assessment content and assessment materials development. Then I briefly discuss the implementation of a portfolio process in which students participated.

Having been heavily influenced by the work on learning environments as social organisations (for example, Waller 1932; Lortie 1975; Rosenholtz 1989), at various points I weave in the more recent ideas of Wenger (1998) regarding communities of practice to highlight the meaning of context in the practice, theory and pedagogy of learning and teaching a course on assessment materials development. In essence, I attempt to tell three stories at once: one about assessment, one about pedagogy and one about trajectories of professional participation.

Narrative inquiry, (second language) teachers and (second language) teacher educators: A brief overview

The last decade of the 20th century saw an interest in narrative inquiry focusing on teachers under various labels, including narrative, story, autobiography, collaborative biography, teachers' curriculum stories, teachers' anecdotes, teacher lore, cognitive perspective, personal knowledge and politics of teaching, among others (see, among others, Connelly and Clandinin 1990; Elbaz 1991; Miller 1992; Schubert and Ayers 1992; Carter 1993; Cochran-Smith and Lytle 1993; Cortazzi 1993; Jalongo and Isenberg 1995; McEwan and Egan 1995; Elbaz-Luwisch 1997; McEwan 1997; Anderson and Herr 1999; Zeichner 1999; Clandinin and Connelly 2000; Gudmundsdottir 2001; Munby, Russell and Martin 2001). These burgeoning academic interests endeavour not only to give voice to teachers but, as Behar-Horenstein and Morgan (1995: 155) contend:

Teacher research makes knowledge about teachers' expertise accessible and provides different perspectives on teaching and learning that can not be known by other forms of inquiry.

In the field of second language education there is also interest in English as a Second Language (ESL) teacher voices and language teacher learning (for example, Bailey and Nunan 1996; Freeman and Richards 1996; Freeman and Johnson 1998; Contreras 2000), professional development and teacher narrative (for example, Johnson and Golombek 2002; Golombek and Johnson 2004), along with stories foreign language teachers tell as they reflect on their professional practice (Hartman 1998a). These stories are important because they:

are bound up in the places, realities, complexities, and personalities that make work interesting, meaningful, and rich for learning. Stripped of these qualities, the substance of their work becomes another bulleted list of abstract points.

(Hartman 1998b: 4)

Surprisingly, there remains a dearth of teacher educator (narrative) inquiry (see, among others, Lanier and Little 1986; Wisniewski and Ducharme 1989; Howey and Zimpher 1990; Ducharme 1993; Judge 1993; Zeichner 1995; Ducharme and Ducharme 1996; Acker 1997). At least since the mid-1980s, calls in the profession to study teacher educators abound. Lanier and Little (1986: 528) suggested there was little known about teacher educators:

But teachers of teachers – what they are like, what they do, what they think – are systematically overlooked in studies of teacher education. Even researchers are not exactly sure of who they are.

Howey and Zimpher (1990: 351) argued that 'studies of the education professoriate need to both better acknowledge historical evolution and attend to the current contexts in which the professors work'. Ducharme and Ducharme (1996: 691) further highlighted the fact that 'Scholars have devoted meagre attention to the serious study of education faculty as a whole, even less to that of teacher education faculty'. In one book that specifically investigated the lives of teacher educators, Ducharme (1993: 2) found that 'Much of the literature on higher education faculty treats education faculty lightly or ignores them totally'. What little is known about teacher educators comes from reviews of what is written in the field. Hazlett (1989), for instance, observed that much of the existing literature about education faculty focused on three genres that included articles concerned with characteristics of the professoriate, articles that sought to define the role, and articles that dealt with a range of conceptual or descriptive matters. (What Thomas [1995] considers mechanistic as opposed to organic and more holistic elements. See Thomas for further discussion.)

More recent forays into teacher educators in the general education literature do provide some updated insights about giving a voice to teacher educators. Russell and Korthagen (1995) show teacher educators reflecting on their personal experiences, and Loughran and Russell (1997) delve into principles and practices which shape teaching about teaching and challenges in teaching and learning about teaching, consider the rethinking of teacher educators' roles and practice, and provide conversations about teacher education. Both volumes help build what Fenstermacher (1997) terms 'scholarly space' for (general) teacher educators' voices.

In the particular area of second language teacher educators, works are also beginning to appear that consider narrative as a means for individual professional development. Casanave and Schecter (1997), for example, offer 19 personal essays of language educators (first- and second-language researchers and practitioners) that variously consider evolving a philosophy,

identity dilemmas, lessons from teachings and learnings, reflections on the professions, and conversations. Their intent is to make connections between the lives of language educators and the themes found in the various essays.

These themes include the weaving together of personal and public strands of lives; the persistent uncertainty of what one knows and the resultant shifting of beliefs over time; the ideological, intellectual and personal conflicts and contradictions within ourselves as well as in our work settings; the ongoing struggle to make meaning in the midst of complexity and confusion; and the enormous diversity of legitimate voices inside and outside the language classroom. (Casanave and Schecter 1997: xxi)

In reviewing Casanave and Schecter's book, Walker (2000: 795) acknowledges unevenness with the essays:

Some authors seem to write for themselves, whereas others imagine an audience of university faculty. Some contributors describe the complex dilemmas facing researchers as theory meets the reality of human experience, whereas others describe L2 learners' experiences with great insight and sensitivity.

However, Walker (2000: 794) contends that there is value with such unevenness because it offers insight into what language education teachers and researchers 'bring to their work', giving 'thought about language, culture, and learning that are personal and even private'.

In other TESOL publications, Bailey et al (1998: 544) present their experiences in collaborative conversations as TESOL (including bilingual and foreign language teachers) teacher educators, where they contend they 'create a dynamic forum in which to negotiate meaning, challenge ideas, and stretch [their] own thinking'. Reflective experiences of TESOL educators are further provided in three case studies in a volume on TESOL teacher education (see Johnson 2000): Johnston (2000) presents a self-study of his dialogical teaching in a Master of Arts Teaching English as a Second Language methods course; Lewis (2000) discusses issues about leading and giving workshops concerning English language teaching in Vietnam; and Guefrachi and Troudi (2000) describe their roles with pre- and in-service professional development courses in the United Arab Emirates. (For a brief overview of narrative research in TESOL, see also Bell [2002] and Pavlenko [2002].)

Clearly, there is still much to do in creating and dealing with teacher educator research as well as teacher educators and narrative inquiry. Ducharme (1993: 11) adamantly maintains that in the wider teacher education field, 'What is needed is an insightful description of faculty lives, their beliefs, and their views of the profession', while in the second language

teacher education literature, Casanave and Schecter (1997: xxi) also call for more work to be completed:

We believe that readers, authors, and editors need to work together to help legitimize the human voice in teaching, research, and study settings, and to push for more serious attention to the human voices and perspective in the language education literature.

Zeichner (1995: 22) promotes similar sentiments but rightfully acknowledges an extremely important and sensitive caveat:

It is time that we paid more attention to the voices of those who do the work of teacher education as we seek to determine the future of the field and to document the wisdom of teacher education practice. However, just as it is dangerous to glorify uncritically the personal and practical knowledge of teachers, it is also dangerous to accept, as necessarily good, everything that teacher educators reveal in their stories about their practice.

The purpose of this paper is to tap the interest in narrative inquiry and apply it to inquiring into a (second language) teacher educator's experience in transforming a postgraduate course. In general, a challenge remains that teacher educators (including ESL teacher educators) are not well researched, narratively or otherwise, within their specific teaching and learning contexts (Goodson 1999), ESL teacher educators even less so. Moreover, the use of narrative offers a slightly different dimension than that used in wider education and second language education literature. Heeding the advice of Zeichner (1995) and others, throughout various sections below I attempt to temper enthusiasm with questions that help demonstrate that the professional development (change) process is continuous and to alert readers and myself to a reflection process in general.

Background

The course discussed here is designed to provide a survey of issues in language testing and assessment to give Master of Arts students experience in critiquing and developing second language (SL) assessment materials. The main text used over the years has been Bachman and Palmer (1996), with other articles and books supplementing the current issues in the SL field (for example, McNamara 1996; Brown 1998), performance assessments (for example, Swanson, Norman and Linn 1995; Delandshere and Petrosky 1998), validity issues (for example, Moss 1994; Messick 1995), and educational assessment and its changes in general (for example, Brookhart 2001; Delandshere 2001, Hargreaves et al 2001). Texts were read each week and lectures, discussions and workshops used throughout the semester.

The bridge between the (theoretical) class discussions and the final (practical) test/assessment product, however, was not well constructed. The final project was to develop a set of assessment materials that took into account the issues tackled in class and which could be used in the particular environments where many of the postgraduate students eventually found themselves as teachers. The typical tasks completed for the course included critiquing three language-testing articles and developing items incorporating Bachman and Palmer's (1996) theoretical framework including their six qualities of usefulness (content validity, reliability, authenticity, practicality, interactiveness and impact). Although student evaluations at semester end were always positive, challenges in getting the students to move from theoretical issues to practical ones often surfaced. Many times, discussions of the readings became difficult in that the students and I would eventually end up selecting particular issues each week from the readings on which to focus, and the wider complex picture of testing was lost in micro-analyses or critiques. Moreover, as there was no one specific test or assessment item to focus on in discussions about theoretical (and practical) issues, it was difficult to relate many of the micro- and macro-issues to assessment materials design, development and implementation.

The practice of assessment development always *followed* the theoretical issues that were read about and discussed in class, and was more time consuming by far. The course needed a change: practice and theory needed to be integrated, and both students and teacher educator needed to be involved in how materials would be used rather than just discussing the issues of materials development. As Wenger (1998: 67) warned:

an excessive emphasis on formalism without corresponding levels of participation, or conversely a neglect of explanations and formal structure, can easily result in an experience of meaninglessness.

To make experiences more meaningful, the following served to help begin building our (supportive) learning community.

Processes in professional development

Ball and Cohen (1999: 13) suggest two important steps for action in professional development:

Centering professional education in practice is not a statement about either a physical locale or some stereotypical professional work. Rather it is a statement about a terrain of action and analysis that is defined first by identifying the central activities of teaching practice and, second, by selecting or creating materials that usefully depict that work and could be selected, represented, or otherwise modified to create opportunities for novice and experienced practitioners to learn.

The first question I needed to address was whether the course focus was going to continue to be on learning the material about language testing and assessment, or to develop language teaching and assessment materials while reading and learning about them within such a process. My gut feeling pointed me in the direction of the latter. Two books particularly reminded me that this latter view was supported by cognitive psychology and designing educational assessment (Bransford, Brown and Cocking 2000; Pellegrino, Chudowsky and Glaser 2001). These readings highlighted the significance of student involvement, while reinforcing the importance of getting students involved in assessment materials construction *and* reading and discussing assessment materials issues. Such literature specifically encouraged me to risk changing my teaching from a content focus to a learner-centred teaching content focus. This idea was not new, but I seemed to have forgotten how to implement it over a period of time within my particular (learning) environment.

The second question I needed to address concerned the selection and/or creation of materials to help practitioners learn. Additional texts from the wider educational testing and assessment literature further reinvigorated and challenged my beliefs and ideas about innovative (language) teaching, testing and assessment (for example, Marzana, Pickering and McTighe 1993; Gipps 1994; Niyogi 1995; Wiggens 1998; Falk 2000; Huba and Freed 2000; Blaz 2001; Stiggins 2001). These texts provided innovative ideas on performance assessment (including the use of the dimensions of learning model), selected response and essay assessment methods, rubric development and use, various types of feedback from students to improve learning, developing standards, assessing reasoning proficiency, skills, and dispositions, and changing thinking from the use of psychometrics to a focus on (second language) educational assessment.

In order to tap students' experiences and the knowledge they bring to their studies (Bransford et al 2000; Pellegrino et al 2001), and to give them the opportunity to create materials that facilitate learning (Ball and Cohen 1999), their first task was to design a test (or assessment items) for a unit of study from a selected textbook. Students were asked to write the test (or assessment items) just as they would if they were teachers in their home learning environment. The task was open-ended in that no specific instructions were given about how to write or present the test (or assessment items) so that students would rely on their experiences as teachers from their particular countries and learning environments. It was also designed to use their personal professional knowledge. Students individually then wrote a test (or assessment items) between the first and second class sessions to

serve as evidence of what they deemed salient in writing assessment materials. The students were asked to focus on one chapter from a particular English language teaching text (Soars and Soars 1993) so they would have material to use in constructing their assessment materials (most, but not all, students were interested in English as a second or foreign language). In essence, these tests became texts for the class to use, to refer to throughout class discussions, and to further develop, change and/or modify as the 13-week semester progressed. It is important to note that these tests remained anonymous, and this anonymity helped to keep arguments or comments focused on the test elements rather than on the test writers per se. At no time did students seem bothered that we were using these tests. In fact, quite a few mentioned the importance and convenience of having the tests available for reflection and discussion.

Focusing on the same chapter proved to be valuable once the tests were written because it demonstrated how students chose items to assess, and provided variations of language constructs (that is, models) that students brought to their test development task. In general, the tests (and assessment items) were focused on grammar items of the text's unit, had various sections that were labelled as testing the four language skills (listening, reading, writing and speaking), and resembled some of the items (or activities) found in the actual textbook. The tests ranged from one typewritten page to eight pages, averaging around three pages. There were multiple-choice questions, fill-in-the-blank and finding mistake items, dialogues and speaking exercises, and paragraph-writing activities, among others.

Having all the tests together gave students the opportunity to initially compare and contrast tests. This highlighted that the students did not really understand some of the instructions of the specific test items and that the majority of test items did not have answer keys to explain or delineate what the test writer was necessarily expecting. At this juncture the students especially seemed to pay attention to the vocabulary and grammar selected for each test. As many of them remarked (and actual comparisons of the tests clearly showed), each example of a test rather randomly had a (test developer's) selection of the issues represented, especially when comparing various vocabulary from the unit (some students aptly recognised that it could end up being a guessing game of sorts for the students completing the test, depending on the particular words selected or not selected). A few began questioning the validity of how and what was selected. The tests provided catalyst for student interactions (with information from the test examples and in discussions with their classmates). Such interactions promoted possibilities to negotiate meaning and reinforced the point that

'Meaning is not pre-existing, but neither is it simply made up. Negotiated meaning is at once both historical and dynamic, contextual and unique' (Wenger 1998: 54). The tests provided us with opportunities to negotiate meaning (individually, among ourselves and with others), develop questions, and reflect variously in dynamic, contextual and unique ways that helped begin addressing the complexity of text construction, implementation and use(s) from their own data sources of tests that they actually created.

It was a vital element to begin including students in these two professional development processes identified from Ball and Cohen (1999) as the course unfolded. Yet questions arise as to when and how to include students, how to develop their need to understand why they need to include them, and the role of teacher and students in the process. How long does it take? Will students be willing to take risks and see the need to complete tasks in a way that may be different from the way they usually participate in classes? How will the teacher respond to the students' various reactions? (These questions are relevant both during and after the course. And the questions are revisited at the start of the next course, too! Answers are not easily attainable.) Such questions also promote the need for further negotiation and meaning-making. In the next section, I tackle the issue of how we began negotiating assessment content in this course.

Negotiating assessment content

During the second class session, students were made aware lectures were not to be a staple product. They would not be told about what they had read in Bachman and Palmer (1996) (or any other text read). Instead Bachman and Palmer would be a resource. The information from the chapters was going to be used and not just discussed. Students were grouped in the second class meeting and asked to consider the main components, strengths and challenges of Bachman and Palmer's six test qualities. (Following Joyce and Weil's (1986) advice on the use of a link-system and ridiculous association to aid memory, these became known as the 'CRAPII' elements from Chapter Two in Bachman and Palmer). Each of the six groups was asked to define one particular quality in their own words, describe the main components, strengths and challenges of the quality, and consider questions about the quality. They then, as a class group, randomly selected and reviewed one of the tests submitted and examined this one test through their one specific quality. As each of the various groups focused on its specific quality and the test, the groups became quite concerned with how little attention was paid to their particular quality and, as a whole group within the class, the students found that most of the six qualities were not

attended to adequately. As one student commented, 'I'm so depressed; I really haven't been testing students' language properly'. Yet we also reminded each other that testing was difficult and that our task was to develop from what we knew and use our experiences and our understandings to begin the improvement process in developing assessment materials.

The groups focusing on the specific qualities were then asked to further consider their component's strengths and challenges after having used it in analysing an actual test. Each group provided an information sheet at the end of the third class concerning each of the six text qualities that was then distributed to all class members the following week. The sheets contained Bachman and Palmer's definition along with other information the various groups discussed.

A group of students defined authenticity as the 'relevance of the test in everyday life and [it] should be related to learners' needs'. They further defined the main components of the quality thus:

We should use authentic materials which are related to learners' needs and can help them to know about the usage of the target language in everyday life. For example, having job interview in oral tests or asking students to write job application letter in writing test.

The strength of the quality was given as connecting 'the classroom to the outside world, so that the students really know what they are doing and why they are doing this'. The challenges included the following information:

It is difficult to create real-life situations within a testing environment. In fact, in many situations this is impossible. It is difficult to even define what is authentic. Not only the test tasks need to be authentic, but also the criteria for marking. It's hard to ask our colleagues to step out of their comfort zone to do what they are not familiar with. School administration may not allow us to have freedom in designing our own tests.

The questions concerning the quality of authenticity included: how do we know what is authentic? How do we define authenticity? From whose perspective do we judge the level of authenticity – the students', the teachers', the test developers', the native speakers'? Individual perspectives and opinions within these groups will also vary. How do we create authentic situations in the classroom? How do the markers account for a test-taker's emotions during the test that would not be present in the real-life situation (for example, being nervous)? The definition of authenticity will also vary for different learning environments, different purposes, different groups of students, difference levels of knowledge etc. How do we take these variations into account?

In the main, the discussion of test qualities kept everyone on task throughout the semester. Which definition do we attempt to implement? As experiences expanded our understandings and dilemmas with the test concepts, how would we make sure that other people would understand our point of view and expanding perspectives? Would our developing knowledge be easily negotiated with students who were going to complete any of the assessment items? Would our understandings and musings be comprehensible? Our discussions took on contextual meaning because we discussed assessment issues from the literature with actual test examples, yet how would/could all this translate into the various learning environments in which teachers eventually found themselves?

The discussions initiated in the class not only challenged knowledges (from both literature *and* students) but also helped create community among the members as they wrestled with the issues from their readings, test examples, experiences and practices (both in the course and outside in various second language classrooms), and worked to develop and create assessment materials together, as will be discussed next.

Negotiating assessment materials development

The students were then asked in the third week to divide themselves into groups of three to work together for the semester to design assessment materials for the same chapter (from Soars and Soars 1993) used for the individual test they had developed between the first and second class sessions. Each week, students worked on developing assessment materials using this chapter. They were asked to assume that they were all working in the same learning environment and teaching the same course so that they might be simulating something that could occur in a real learning environment.

It became clear to course members during this group work that just writing an item (as they did in writing their initial tests) was far from sufficient. For these students, the time spent defining constructs, developing and piloting assessment materials, and rewriting and rethinking the various assessment tasks and items was quite burdensome. Although the importance of such actions in assessment materials development is noted in the assessment literature (for example, Bachman and Palmer 1996; Wiggins 1998; Stiggins 2001), we found the actual process cumbersome and not part of our typical repertoires as teachers in particular learning environments (as many students noted in comments to others in class and me and also in reflections from their portfolios that they usually just wrote a test like that written at the beginning of the semester – and their initial tests served as additional evidence to support such a contention). We found we needed to develop, pilot, consider the responses we received, and then reconsider the

particular assessment item. We found that we might have to start with a construct, or discuss topical knowledge, or create a rubric before we even considered the type of assessment or test item.

Our group experiences in developing assessment materials were not a linear process (Bachman and Palmer 1996; Davidson and Lynch 2002): issues had to be discussed and revisited. Moreover, we were participating in what Davidson and Lynch described as the power of group process and testcrafting. Only, instead of being part of a group in a (selected) team, we were part of a university course developing and learning about assessment materials. It was during these various activities and perspectives that the ideas we were relying on from readings supported and reinforced the ideas put forth by Davidson and Lynch. It was also evident that this helped establish what it meant for us to be a competent participant in this community (Wenger 1998) as we developed, defined and worked with assessment materials. In the first few weeks, one group developed an item for high school-aged students in a traditional linear sense in that they relied upon themselves for determining what and how a certain area should be assessed. They felt it was too much trouble to think about all the 'theoretical' issues. They did agree to pilot the item and one of the group's members gave it to his 14-year-old daughter (a second language user of English) to attempt to complete. The teenager told her father in no uncertain terms that she did not understand the instructions that well and that, as far as she could understand it, the activity would be a waste of her time. Her father returned to the group with this news. The traditional, typical linear process of this particular group's members was shattered and the group members started seeing the meaning of test design as craft; they started the process again with a different vision.

We further found that we continued revisiting the six qualities of testing and added varying perspectives, questions and challenges encountered throughout our experiences. The test qualities acquired extended (and varying) meanings and intentions throughout the semester-long course. For example, did reliability deal with the material chosen or the reliability of the students participating, or both (and perhaps anything else)? Reliability changed to mean more that just concern for people marking assessments. Practicality also took on various understandings. Whose practicality was the group interested in pursuing – the test developers', the students', the teachers'? Would the results be 'practical' enough for parents, administrators and other teachers to understand? As groups considered and more thoroughly delineated the impact of the tests they were working with from various shareholders' perspectives, their notions of impact became multi-dimensional, and they moved from a focus just on the interaction between

the individual and the task to a focus also on interacting with others and text. Thus, interactiveness took on meaning additional to that described initially by Bachman and Palmer. Similarly the quality of construct validity created further inquiry. Bachman and Palmer (1996: 21) define it as 'the meaningfulness and appropriateness of the interpretations that we make on the basis of test scores'. Interestingly, the 'we' pronoun became something to be investigated: just whose interpretations were being made, how were they being made, and why were they being made in such ways? More questions than answers transpired from interpreting construct validity. One student highlighted the fact that topical knowledge – as Bachman and Palmer (1996: 65 write, 'What we will call topical knowledge [sometimes referred to as knowledge schemata or real-world knowledge] can be loosely thought of as knowledge structures in long-term memory') – was not something that could just be described but that needed to include what students brought to language classrooms and how teachers and students would have to negotiate topical knowledge in their daily classroom activities.

As we interacted with text, talk and people, we were all questioning the amount of time it actually took to prepare what we were developing as the terms for adequate assessment materials development. Our queries included how to complete such work within the day-to-day tasks of being within a learning environment, when to find time to seek the assistance of our colleagues, and preparing enough in advance to begin involving students in trial processes. Further, we asked more questions about the six qualities: can all six qualities be evidenced in one particular assessment item? What if one quality becomes more dominant than another? Could we consider two or three qualities in one assessment item and then consider two or three others in another? What would this mean to assessment materials development within our particular learning environments? Interestingly, we also came to rely on each other within various groups, and within the course in general. How would we find the potential of such collegial and collaborative ethos out in the real world? Our questions filled us with both trepidation and excitement.

Such interactive professional practice reminded us of what Wenger (1998: 54) might have meant when he wrote:

meaning is always the product of its negotiation, by which I mean that it exists in the process of negotiation. Meaning exists neither in us, nor in the world, but in the dynamic relation of living in the world.

Many of us participating in the various activities were amazed that the definitions of the qualities from the book failed to capture the nuance, subtleties and multiple meanings the qualities manifested in actually

developing assessment materials. Our community was interpreting, expressing and negotiating meaning about the practices and theories related to assessment materials while experiencing the non-linear process of assessment materials development.

Negotiating further assessment content

Together, we read Shohamy's (2001) critical perspective on language tests. This treatise offered background into why the SL profession should look more carefully at test development and challenged SL teachers to become more aware of how and why they use SL tests and assessments in their classrooms. The book developed a case that SL testing needed attention on many fronts. The students found Shohamy's book enthralling, with the majority commenting on how she nicely outlined many problems they had previously thought about but had not cohesively put together as well as she had. Additionally, Alderson and Banerjee (2001; 2002) highlighted the increasingly complex SL testing field by reviewing topics found in the literature including, among others, washback, ethics in language testing, politics, standards in testing, computer-based testing and challenges for the future. These texts offered information which complemented that found in Bachman and Palmer (1996), and together all four texts formed the required reading texts for the course. During the period from week three to week eight, we (briefly – 30 to 40 minutes) considered and discussed issues from Shohamy, and Alderson and Banerjee. We also allowed time for group work to continue creating assessment materials. Weeks eight to twelve were used mainly for continuous group work. It was during this time that we reminded each other about the six test qualities, revisited issues in previous readings, shared with each other readings that were done outside the class, and shared with each other challenges in the development of assessment materials. During week 13 the five groups shared their test items in 20-minute presentations that allowed for some questions and further discussion.

In creating our assessment materials as we discussed various testing content, we became aware of the growing amount of literature in the field. How do we find time to read, digest and implement the various ideas? How do we select readings from the literature? What will help us complete our task quicker? These content issues and activities promoted further opportunities for negotiating meaning. As Wenger (1998: 262) notes:

Learning from this perspective is a very dynamic and systemic process in which mutual alignment continually plays the role of catalyst. This focus on the negotiation of meaning is a focus on the potential for new meanings embedded in an organization.

In our case meaning making was also a focus on the potential for new meanings embedded in the extension and development of our communal knowledge. That is, we developed meaning and understanding of issues in relation to how we talked with each other, how we interacted with each other, with the various texts read, and with the various materials we were developing. We were participating in a classroom where learning was seen as a potential social system productive of possible new meanings (Wenger 1998) that could serve as an example of what type of learning environment we could engender in each of our own classrooms elsewhere. Keeping track of such learning experiences for later use was challenging, but the use of portfolio development assisted with such a task and that activity will be discussed next.

Portfolio: Designing, negotiating and documenting individual learning

Johnson and Johnson (1991) consider individual accountability as one of the five basic elements for lessons to be cooperative and help scaffold students' learning (professional development) and to assist students eventually in developing an individual test/assessment project. A portfolio documenting their learning and professional development was therefore required of all students. Also, having read Shohamy (2001) in preparation for the class (and eventually as a class text), I was aware that I should try to include students' voices and ideas. With this in mind, I was much less directive in this particular assessment activity. This caused much concern throughout the semester, as I purposely gave no specific guidelines about what students needed to put in their portfolios. Nonetheless, I put together a four-page overview of portfolio issues and shared it with them, including an overarching definition of portfolio assessment (Huba and Freed 2000: 233):

In short, it is the purpose of portfolio assessment to help the learners become integral and conscious participants in their learning processes, by having them recognize both individual responsibility and ownership within that process, and by having them become interactive partners with the teacher in shaping that learning process.

Ideas and narrative excerpts concerning portfolios such as gathering and giving feedback, some essentials of portfolios (for example, systematic collections, values that underlie the use of portfolios and product elements), classroom assessment techniques (for example, the minute paper, muddiest point, word journal) and numerous reflection questions were taken from various sources (Gaudiani 1981; Valencia 1990; Tierney,

Carter and Desai 1991; Angelo and Cross 1993; Roe and Vukelich 1994; Brookfield 1995; Huba and Freed 2000). I also asked that students focus on their learning and what they did during the course so that they did not necessarily have to do any extra work to make items or entries for the portfolio. They were encouraged to develop their own ways to document or demonstrate in their portfolios how they used the various discussions, readings and group work from the course. Because of initial apprehensions among the students, I asked them to complete certain short activities during the first five or six classes to encourage various ways of documenting their work, and, after requests from students, I offered to periodically read and make written feedback on their portfolio entries throughout the semester. In order to help settle some student unrest, I gave a more generic marking rubric to help them consider the types of things to put in their portfolios (this rubric was used in my other classes as a general marking rubric for essays and activities and I thought it might help the students better focus their entries by reminding them about providing evidence that served as evidence to support their developing comprehension, application, analysis and synthesis skills).

At one time early in the semester during a class discussion, one student took me to task and told me in no uncertain terms that I placed students at a disadvantage because I provided neither an example of a portfolio nor told the students exactly what to do in compiling their portfolios. Other students shared this particular student's feelings. I asked that all continue developing their portfolios in their various ways and to make sure that they handed in entries for comments. I shared that I was trying to heed Shohamy's (2001) advice but this did not seem to offer any consolation (even though it allowed us to discuss the issues and offer an option for more reflection). I responded quite positively in all cases to the items received from the students during the semester as they were all developing *more than adequate* entries to their portfolios, in my opinion (and I shared this with them). I reassured them weekly that all would be well and to try not to worry. The portfolio was due the last day of class, along with a one-to two-page narrative that reflected on the portfolio and a final mark they would give themselves for the portfolio. The students did not know until the final class session (once all portfolios were handed in) that they were going to receive the actual mark they each had given themselves for this portion of the course assessment. (I had kept this a secret, as I was afraid if they had actually known this, they would not have taken the task of reflection seriously and implemented criteria on the assessment of their own work.) In order to include the voices of those participating in the assessment, I was attempting here to relinquish my power and include

the emerging power of student input into the assessment process (see discussions in Shohamy 2001). I might have been expecting too much, too soon, as when they found out that the mark they had given themselves would be the mark calculated within their final mark, some were disappointed that I would not be assigning them a mark for this. Some seemed to want me to retain my power in the process and leave me with the task of determining the final mark. In conferences with each student during the semester examination period, we discussed the criteria they used for their particular mark, how to look for evidence within their portfolio entries and the portfolio process in general. Some students seemed to have difficulty building a relationship between these issues, while others felt that the general criteria within the information and rubrics they were given throughout the semester helped them focus on what type of entries to include in the portfolio. I realised some students did not want me to give up power; some students did not want to gain it. I did feel, though, that the students fairly and quite rationally gave themselves appropriate marks with the data they accumulated and the manner in which they evaluated themselves. I learned that I was not only trying to empower them, but I was facilitating them to understand the relationship between evidence and marking – planting the seeds for empowerment. (I would not do this a second time, as future students might be told by these students about the marking. Therefore, I now inform students at the beginning of the semester that they will mark their portfolio and the final mark for the portfolio will then be negotiated between each individual student and me, with the student taking the lead on defending their mark rationale – so far only one student and I have disagreed and a compromise in the mark was then agreed upon.) Weimer (2002) discusses in detail the difficulties of developing self-assessment issues; I find the dilemmas discussed by Weimer are also the many dilemmas I face in my teaching (regardless of subject matter). The portfolio exercise was controversial; nonetheless, I find that it is the controversy that pushed my students and me to further negotiate and understand the issues. Not only have I continued to be vague about the portfolio process at the beginning of the subsequent assessment course, I have added the portfolio to another class concerning issues in program development.

Regardless of the ups and downs of this portfolio experience, students found that the process of compiling the portfolio gave them useful insights and was important to their development. As two students noted in their end-of-semester evaluations:

Self assessment is new to me. Portfolio assessment is new to me. Nevertheless, when I look at what I have included in my portfolio, I can see that I have learned quite a bit about testing and assessment.

and

The new approach toward portfolio assessment is really useful to assess actual students in real settings. We actually experienced the difficulties and challenges, and limitations of teaching and assessment, which was good.

A number of students have since allowed me to photocopy their portfolios and have them placed in the library high-use area for future students to peruse. The individual portfolio activities allowed an additional avenue for community building (Wenger 1998), providing evidence that learners can be involved in the design of their own learning. The students actually decided what they needed to learn by including elements in their portfolios that demonstrated such a process. They participated more in their learning process rather than relying on the teacher to provide the learning process for them; initially most were quite hesitant, but I tried to guide them through activities and encouraged them to rely on me less and less throughout the semester.

The portfolio process continues to be the most challenging part of this course. Trying to help the students move from relying solely on the teacher to relying on themselves is most challenging. Reflections on past experiences show, however, that each time it is not just a new experience for the students, but a new one for me. How do I empower students who might not be open to empowerment? How do I negotiate my intentions and students' intentions? How can I become more ethically aware of the processes I use and then how can I become more nurturing with the decisions that are arrived at or developed in the particular course, activity and/or event? The reflection process on portfolios never ceases.

Discussion

In this particular manuscript I developed a narrative about professional development in a university setting for a course focusing on second language assessment and testing, taking into account pedagogical and content issues in the professional development of second language teachers and second language teacher educators (see Glesne and Peshkin 1992 regarding the importance of audience in creating and presenting text[s]). The most exciting part of the whole process was learning and professionally developing at the same time; learning as professional development, and not learning as a by-product of professional development (ie, learning through

action and reflection and not learning through being told something or to do something). My students and I got to share our enthusiasm of learning with each other (and I now get to share our experiences with second language teachers and teacher educators in this article), rather than simply learning a certain amount of material and implementing that information. I personally learned what Wenger (1998: 134) meant when he wrote:

Communities of practice should not be reduced to purely instrumental purposes. They are about knowing, but also about being together, living meaningfully, developing a satisfying identity, and altogether being human.

Towards the end of the semester, two unexpected experiences from students highlighted some additional learning (professional development) insights. The first one occurred when a student came to class and shared a test she had given in her particular learning community that day. She started by saying that usually she just gave the test to the students. She did have the students take the test, but while they were completing it, she looked at it again, in the light of the CRAPII elements we had been discussing on the course. She was surprised how poorly conceptualised and written the test was, and how test quality issues she and the class were learning had not been taken into account: she wanted to bring it to the attention of the class as an example of a poor test used in a real classroom. She also related how she now looked at tests quite differently since participating in the class. Her classmates readily agreed with her after perusing copies of the test.

The second experience was related to the class on the last day by two teachers in the class who were teachers who worked in the same teaching community, and who had started people asking what they were doing when they were piloting some of the items in their classrooms. A group of five to seven teachers started working together in that particular context, with assessment materials and teaching activities as well. The two students said they followed many of the ideas and activities they tackled in this class.

What these two examples show is that the focus on learning in practice did affect students in ways that I had not experienced in previous years. Moreover, our negotiations and practices within our particular course gave contextual meaning to what Shulman (1999: 12) meant when he wrote:

Learning flourishes when we take what we think we know and offer it as community property among fellow learners so that it can be tested, examined, challenged, and improved before we internalize it.

Or as Wenger (1998: 52) contends, '*Practice is about meaning as an experience of everyday life*' (my italics).

In course evaluations at the end of the semester in which I first taught the course in this manner, students related their feelings about the class

and about me as their teacher, which helps further clarify how they felt about being part of this type of learning community. Among others, students wrote:

Excited, contemporary approach with very up-to-date ideas.

I would like to say that I enjoyed this course very much because I was deeply involved in the course by good selection of readings, multiple assessment methods and personal communication with the lecturer in consultation time.

Application of what we have learned and discussed in the class. This is the best class that I have since my first and second semester.

The strengths are helping me to understand more about language testing. It is very helpful and I think it is a must [that] all teacher[s] should know and keep up with it. It is very good to have a workshop, especially with people from different countries. It's good to learn from each other, and it really help[s] me to see things more openly, widely. When we learn how to teach the language, it's no use not to learn how to test. Then, this course is the major component of being a language practitioner, it's good [second language user of English].

The material was presented in such a way that it encouraged students to take that information and apply it to a practical setting. The course also gave [me] a new-found awareness of different elements and considerations in the field of testing.

The teacher included the students in the process of assessment and therefore gave students more control of their learning process.

The lecturer's enthusiasm and passion for the subject has a great impact on the atmosphere, which inspires me to study; the lecturer's good selection of materials (readings) for students; he always tries to help students through. It's excellent to have opportunity to have a say. Perfect knowledge of the subject; enthusiastic in teaching; he brings new things to think about all the time.

(The quantitative descriptive statistics of the final evaluation questionnaire also indicated that students appreciated the course, and my academic supervisor acknowledged that he had not seen such impressive end-of-course evaluation results.)

It is important to briefly highlight that the particular approach to professional development taken in this course was different to the style of teaching and learning found in typical postgraduate classes at this university, and this was something that the students and I struggled with most of the semester. To say the least, there were uncomfortable moments. As one student suggested at the end of the semester when asked to comment on how to improve the course, 'Students should be made to understand [the] lecturer's teaching philosophy more so to avoid possible misunderstanding'. There was also apprehension from many students

because the teacher seemed to be less in charge and more interested in joining in the discussion. As I remember one student commenting to me, 'You don't tell us what to do, you just ask many questions to make us think. You at least though try to begin answering them, too.' Instead of reading texts, discussing texts, participating in activities about texts, and then developing assessment, we developed assessment materials while reading and discussing texts and participated in activities that encouraged the development of assessment materials. We continually reminded ourselves that participation in learning was in essence the act of professionally developing our acumen while creating testing and assessment items:

It [the course] really helped me to understand testing and its components in regard to the CRAPII elements. By doing rather than just being lectured on how to test, I have really started to understand what testing really involves. I feel confident that I could help develop useful texts in the real world. (I really enjoyed the course, thanks).

In essence, the students and I put into action Ball and Cohen's (1999) ideas to centre learning (professional development) in practice.

Naturally, various class members suggested improvements, including more time for group work, more individual instruction, and not such a heavy workload. Additionally, they mentioned the need to be clearer about assessment tasks, to improve and make clear/explicit the criteria for portfolio assessment, and to supplement the readings of Shohamy (2001) and Bachman and Palmer (1996). One person wrote:

Perhaps a little more structure in assessment items is needed for some students but personally I felt the openness allowed me to have more of a role in my learning.

And another warned that, although excited about implementing something new, 'it's not the case to take risks for certain students'. These issues remind me that one of my tasks as a second language teacher educator is to continue to grow, change, and be challenged. In making course outlines and developing activities and assessment items for future courses, my reflections, experiences, what I have read and students' comments continue to influence me in a number of ways: they may guide, sway, encourage and motivate me, or cause me to question, interrogate, raise doubt, explain or concede, and they may also rankle at times. Perhaps it is the conundrums of being a (second language) teacher educator that keep me interested in challenging teacher practitioners to become practical and theoretical assessment materials developers within a developing, supportive, professional community.

REFERENCES

- Acker, S. (1997). Becoming a teacher educator: Voices of women academics in Canadian faculties of education. *Teaching & Teacher Education*, 13(1), 65–74.
- Alderson, J. C., & Banerjee, J. (2001). Language testing and assessment (Part 1). *Language Teaching*, 34, 213–236.
- Alderson, J. C., & Banerjee, J. (2002). Language testing and assessment (Part 2). *Language Teaching*, 35, 79–113.
- Anderson, G. L., & Herr, K. (1999). The new paradigm wars: Is there room for rigorous practitioner knowledge in schools and universities? *Educational Researcher*, 28(5), 12–21, 40.
- Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers*. San Francisco, CA: Jossey-Bass.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bailey, F., Hawkins, M., Irujo, S., Larsen-Freeman, D., Rintell, E., & Willett, J. (1998). Language teacher educators collaborative conversations. *TESOL Quarterly*, 32(3), 536–546.
- Bailey, K. M., & Nunan, D. (Eds.). (1996). *Voices from the language classroom*. Cambridge: Cambridge University Press.
- Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 3–32). San Francisco, CA: Jossey-Bass.
- Behar-Horenstein, L. S., & Morgan, R. R. (1995). Narrative research, teaching and teacher thinking: Perspectives and possibilities. *Peabody Journal of Education*, 70(2), 139–161.
- Bell, J. S. (2002). Narrative inquiry: More than just telling stories. *TESOL Quarterly*, 36(2), 207–213.
- Blaz, D. (2001). *A collection of performance tasks and rubrics foreign languages*. Larchmont, NY: Eyes on Education.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.
- Brookfield, S. (1995). *Becoming a critically reflective teacher*. San Francisco, CA: Jossey-Bass.
- Brookhart, S. M. (2001). Successful students' formative and summative uses of assessment information. *Assessment in Education*, 8(2), 153–169.
- Brown, J. D. (Ed.). (1998). *New ways of classroom assessment*. Alexandria, VA: TESOL.
- Carter, K. (1993). The place of story in the study of teaching and teacher education. *Educational Researcher*, 22(1), 5–12, 18.
- Casanave, C. P., & Schecter, S. R. (Eds.). (1997). *On becoming a language educator: Personal essays on professional development*. Mahwah, NJ: Lawrence Erlbaum.

- Clandinin, D. J., & Connelly, F. M. (2000). *Narrative inquiry: Experience and story in qualitative research*. San Francisco, CA: Jossey-Bass.
- Cochran-Smith, M., & Lytle, S. L. (1993). *Inside outside: Teacher research and knowledge*. New York, NY: Teachers College Press.
- Connelly, F. M., & Clandinin, D. J. (1990). Stories of experience and narrative inquiry. *Educational Researcher*, 19(5), 2–14.
- Contreras, G. A. (2000). Self-storying, self-understanding: Toward a narrative approach to EFL teacher education. *TESOL Journal*, 9(3), 24–27.
- Cortazzi, M. (1993). *Narrative analysis*. Washington, DC: Falmer Press.
- Davidson, F., & Lynch, B. K. (2002). *Testcraft: A teacher's guide to writing and using language test specifications*. New Haven, CT: Yale University Press.
- Delandshere, G. (2001). Implicit theories, unexamined assumptions and the status quo of educational assessment. *Assessment in Education*, 8(2), 113–133.
- Delandshere, G., & Petrosky, A. R. (1998). Assessment of complex performances: Limitations of key measurement assumptions. *Educational Researcher*, 27(2), 14–24.
- Ducharme, E. R. (1993). *The lives of teacher educators*. New York, NY: Teachers College Press.
- Ducharme, E. R., & Ducharme, M. K. (1996). Development of the teacher education professoriate. In F. B. Murray (Ed.), *The teacher educator's handbook* (pp. 691–714). San Francisco, CA: Jossey-Bass.
- Elbaz, F. (1991). Research on teachers' knowledge: The evolution of a discourse. *Journal of Curriculum Studies*, 23(1), 1–19.
- Elbaz-Luwisch, F. (1997). Narrative research. Political issues and implications. *Teaching & Teacher Education*, 13(1), 75–83.
- Falk, B. (2000). *The heart of the matter: Using standards and assessment to learn*. Portsmouth, NH: Heinemann.
- Fenstermacher, G. D. (1997). Foreword. In J. Loughran & T. Russell (Eds.), *Teaching about teaching: Purpose, passion and pedagogy in teacher education* (pp. vii–xiii). Washington, DC: Falmer Press.
- Freeman, D., & Johnson, K. E. (1998). Reconceptualizing the knowledge-base of language teacher education. *TESOL Quarterly*, 32(3), 397–417.
- Freeman, D., & Richards, J. C. (Eds.). (1996). *Teacher learning in language teaching*. Cambridge: Cambridge University Press.
- Gaudiani, C. (1981). The participatory classroom. *Foreign Language Annals*, 14(3), 173–180.
- Gipps, C. V. (1994). *Beyond testing: Towards a theory of educational assessment*. London: Falmer Press.
- Glesne, C., & Peshkin, A. (1992). *Becoming qualitative researchers: An introduction*. New York, NY: Longman.

- 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(3), 307–327.
- Goodson, I. F. (1999). Representing teachers. In M. Hammersley (Ed.), *Researching school experience: Ethnographic studies of teaching and learning* (pp. 122–133). London: Falmer Press.
- Gudmundsdottir, S. (2001). Narrative research on school practice. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed.) (pp. 226–240). Washington, DC: American Educational Research Association.
- Guefrachi, H., & Troudi, S. (2000). Enhancing English language teaching in the United Arab Emirates. In K. E. Johnson (Ed.), *Teacher education* (pp. 189–204). Alexandria, VA: TESOL.
- Hargreaves, A., Earl, L., Moore, S., & Manning, S. (2001). *Learning to change: Teaching beyond subjects and standards*. San Francisco, CA: Jossey-Bass.
- Hartman, D. K. (Ed.). (1998a). *Stories teachers tell: Reflecting on professional practice*. Lincolnwood, IL: National Textbook Company.
- Hartman, D. K. (1998b). The sea of stories. In D. K. Hartman (Ed.), *Stories teachers tell: Reflecting on professional practice* (pp. 3–13). Lincolnwood, IL: National Textbook Company.
- Hazlett, J. (1989). Education professors: The centennial of an identity crisis. In R. Wisniewski & E. Ducharme (Eds.), *The professors of teaching: An inquiry* (pp. 110–128). Albany, NY: State University of New York Press.
- Howey, K. R., & Zimpher, N. L. (1990). Professors and deans of education. In W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 349–370). New York, NY: Macmillan.
- Huba, M. E., & Freed, J. E. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning*. Boston, MA: Allyn and Bacon.
- Jalongo, M. R., & Isenberg, J. P. (1995). *Teachers' stories: From personal narrative to professional insight*. San Francisco, CA: Jossey-Bass.
- Johnson, D. W., & Johnson, R. (1991). *Learning together and alone: Cooperative, competitive, and individualistic learning* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Johnson, K. E. (2000). *Teacher education*. Alexandria, VA: TESOL.
- Johnson, K. E., & Golombek, P. R. (Eds.). (2002). *Teachers' narrative inquiry as professional development*. Cambridge: Cambridge University Press.
- Johnston, B. (2000). Investigating dialogue in language teacher education: The teacher educator as learner. In K. E. Johnson (Ed.), *Teacher education* (pp. 157–173). Alexandria, VA: TESOL.
- Joyce, B., & Weil, M. (1986). *Models of teaching* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

- Judge, H. (1993). Foreword. In E. R. Ducharme, *The lives of teacher educators* (pp. vii–ix). New York, NY: Teachers College Press.
- Lanier, J. E., & Little, J. W. (1986). Research on teacher education. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.) (pp. 527–569). New York, NY: Macmillan.
- Lewis, M. (2000). Lessons from a jet in jet out expert: Cooperation, adaptability, and relevance in Vietnam. In K. E. Johnson (Ed.), *Teacher education* (pp. 175–188). Alexandria, VA: TESOL.
- Lortie, D. C. (1975). *Schoolteacher: A sociological study*. Chicago, IL: University of Chicago Press.
- Loughran, J., & Russell, T. (1997). *Teaching about teaching: Purpose, passion and pedagogy in teacher education*. Washington, DC: Falmer Press.
- Marzana, R. J., Pickering, D., & McTighe, J. (1993). *Assessing student outcomes: Performance assessment using the dimensions of learning model*. Alexandria, VA: Association for Supervision and Curriculum Development.
- McEwan, H. (1997). The functions of narratives and research on teaching. *Teaching & Teacher Education*, 13(1), 85–92.
- McEwan, H., & Egan, K. (Eds.). (1995). *Narrative in teaching, learning, and research*. New York, NY: Teachers College Press.
- McNamara, T. (1996). *Measuring second language performance*. New York, NY: Longman.
- Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice*, 14, 5–8.
- Miller, J. L. (1992). Teachers' spaces: A personal evolution of teacher lore. In W. H. Schubert & W. C. Ayers (Eds.), *Teacher lore: Learning from our own experience* (pp. 11–24). New York, NY: Teachers College Press.
- Moss, P. A. (1994). Can there be validity without reliability? *Educational Researcher*, 23(2), 5–12.
- Munby, H., Russell, T., & Martin, A. K. (2001). Teachers' knowledge and how it develops. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed.) (pp. 877–904). Washington, DC: American Educational Research Association.
- Niyogi, N. S. (1995). *The intersection of instruction and assessment: The classroom*. Princeton, NJ: Educational Testing Service.
- Pavlenko, A. (2002). Narrative study: Whose story is it, anyway? *TESOL Quarterly*, 36(2), 213–218.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- Roe, M., & Vukelich, C. (1994). Portfolio implementation: What about r for realistic? *Journal of Research in Childhood Education*, 9(1), 7–14.
- Rosenholtz, S. J. (1989). *Teachers' workplace: The social organization of schools*. New York, NY: Longman.

- Russell, T., & Korthagen, F. (Eds.). (1995). *Teachers who teach teachers*. Washington, DC: Falmer Press.
- Schubert, W. H., & Ayers, W. C. (Eds.). (1992). *Teacher lore: Learning from our own experience*. New York, NY: Teachers College Press.
- Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. New York, NY: Longman Pearson Education.
- Shulman, L. S. (1999). Taking learning seriously. *Change*, July/August, 11–17.
- Soars, L., & Soars, J. (1993). *Headway* (Student's book, Elementary). Oxford: Oxford University Press.
- Stiggins, R. J. (2001). *Student involved classroom assessment* (3rd ed.). Prentice Hall College Division.
- Swanson, D. B., Norman, G. R., & Linn, R. L. (1995). Performance-based assessment: Lessons from the health professions. *Educational Researcher*, 24(5), 5–11, 35.
- Thomas, D. (1995). Treasonable or trustworthy text: Reflections on teacher narrative studies. In D. Thomas (Ed.), *Teachers' stories* (pp. 1–23). Philadelphia, PA: Open University Press.
- Tierney, R. J., Carter, M. A., & Desai, L. E. (1991). *Portfolio assessment in the reading-writing classroom*. Norwood, MA: Christopher Gordon Publishers, Inc.
- Valencia, S. (1990). A portfolio approach to classroom reading assessment: The whys, whats, and hows. *The Reading Teacher*, 43(4), 338–340.
- Walker, C. L. (2000). Review of Casanave & Schecter's *On becoming a language educator: Personal essays on professional development*. *TESOL Quarterly*, 34(4), 794–795.
- Waller, W. (1932). *The sociology of teaching*. New York, NY: John Wiley & Sons, Inc.
- Weimer, M. (2002). *Learner-centered teaching*. San Francisco, CA: Jossey-Bass.
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge: Cambridge University Press.
- Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco, CA: Jossey-Bass.
- Wisniewski, R., & Ducharme, E. R. (1989). Why study the education professoriate? An introduction. In R. Wisniewski & E. R. Ducharme (Eds.), *The professors of teaching: An inquiry* (pp. 1–9). Albany, NY: State University of New York Press.
- Zeichner, K. M. (1995). Reflections of a teacher educator working for social change. In T. Russell & F. Korthagen (Eds.), *Teachers who teach teachers* (pp. 11–24). London: Falmer Press.
- Zeichner, K. M. (1999). The new scholarship in teacher education. *Educational Researcher*, 28(9), 4–15.