



This is the post-print of Shay, S. & Ocklers, L. 2009. Judging essays: the challenge of interpretation. *South African Journal of Higher Education*. 23(1): 142-154.

It is made available according to the terms of agreement between the author and the journal, and in accordance with UCT's open access policy available:

<http://www.openuct.uct.ac.za/sites/default/files/UCTOpenAccessPolicy.pdf>, for the purposes of research, teaching and private study.

# Judging essays: The challenge of interpretation

L. Olckers

School of Public Health and Family Medicine  
University of Cape Town  
South Africa  
e-mail: lorna.olckers@uct.ac.za

S. Shay

Centre for Higher Education Development  
University of Cape Town  
South Africa

## Abstract

In the past decade there has been increasing interest in assessment as a social practice. There is a growing recognition that assessment, particularly of complex tasks, is a judgment which is socially-situated and thus contingent on a variety of factors which constitute the assessment event. Drawing on the hermeneutical tradition, this article explores the interpretive processes of essay markers in two Health Science courses at the University of Cape Town, with a particular focus on markers' judgements about language. The analysis exposes multiple 'circles' of influence on markers' professional judgements, including the texts (both the written product and the student writer), the marker's interpretive framework, the interpretive community and the wider institutional discourses. The article contributes to a better understanding of the crucial role of the interpretive community in the validation of assessment interpretations.

## INTRODUCTION

In the past decade there has been increasing interest in assessment as a social practice (Gipps 1999; Filer 2000; Broadfoot 1996). In contrast to traditional notions of assessment as an accurate, reliable, scientific process for measuring student achievement, there is growing recognition that assessment, particularly of complex tasks, is a judgment which is socially-situated and thus contingent on a variety of factors which constitute the assessment event – the learning context, the nature of the assessment task, the purposes of the assessment and the relationship between the assessor and the assessed. Thus the central argument is that the assessment of complex tasks is a socially-situated interpretive act (Shay 2004, 2005).

Assessment as an interpretive process has important implications for validity. With the shift to assessment as interpretation, scholars argue that validity is best understood as the extent to which interpretations are sound for their intended purposes (Messick 1989; Nitko 2001). This understanding suggests that validity is

not a fixed property of a particular assessment task or set of results, but is always contingent upon the interpretations and purposes for which those interpretations are used. Given this emphasis on validity as an evolving rather than a fixed property, some validity scholars suggest that the preferable term is 'validation' as it signals validity as an on-going process of strengthening the alignment between assessment interpretations and the educational purposes they serve (Cronbach 1989).

There have been a number of studies which give some insight into the judgment-making processes of higher education teachers in the context of a rapidly diversifying higher education system – diversity of students as well as staff (Wyatt-Smith and Castleton 2005; Reed, Granville, Janks, Makoe, Stein, Van Zyl and Samuel 2003; Shay 2003). This article contributes to a better understanding of these judgment-making processes, with a particular focus on markers' judgements about language. Drawing on the hermeneutical tradition, it explores markers' interpretive processes within the context of two Health Science courses at the University of Cape Town. The conclusion returns to the implications for the validation of these assessment interpretations.

## **BACKGROUND**

The Health Sciences Faculty at the University of Cape Town has undergone significant change since the 2002 adoption of a 'new' or 'revised' MB.Ch.B. curriculum. This shift has resulted in a rigorous process of review and redesign of educational outcomes and activities consistent with problem-based learning theory. Ensuring the alignment between outcomes, activities and assessment tasks, as well as consistency in marking, has been central to the curriculum review process.

The courses 'Becoming a Professional' (BP) and 'Becoming a Health Professional' (BHP) form the core curriculum for all first year Health Sciences students at the University of Cape Town. The approximately three hundred and fifty BP and BHP students are divided into groups of twelve, each group lead by a facilitator. Learning in BP/BHP takes place largely through these facilitator-led small groups where activities are of a mostly experiential nature. These activities are supported by prescribed readings, reflective portfolio tasks and essay assignments all contributing to the broad outcome of developing students into knowledgeable, empathic and reflective 'integrated health professionals' (Olckers et al. 2006). The essays assist students with the consolidation of knowledge and the development of writing skills linked to their future professional roles. Essay topics include themes of listening and other interpersonal skills, health and human rights and the South African health system.

The BP/BHP small group facilitators are graduates from a variety of disciplines, including social work, psychology, teaching, nursing, physiotherapy and law. They are selected on the basis of their enthusiasm for working with first year students, experience in working with small groups, openness to more experiential methods of teaching and a willingness to learn. Participation in intensive training is compulsory and on-going, and through this activity facilitators form a tight and supportive team

which encourages an atmosphere of empathy which is reflected in the pedagogical ethos of the courses.<sup>1</sup>

Significant challenges face the teaching staff given the diversity in educational preparedness of the students. This diversity continues to reflect the racialised divisions of Apartheid policy when the provision of state schooling was divided according to South Africa's various 'race' groups (i.e. White, Indian, Coloured, and African), with each resourced in descending order of priority by separate departments of education. The result was severely under-resourced schooling for both African and Coloured students.

Compounding the poor quality of educational provision is the issue of language. Despite the national policy on mother tongue medium of instruction, many African and Coloured children are instructed in a language other than their home language. The result is that these students have neither a sufficient conceptual foundation in their home language nor sufficient academic proficiency in English (Kapp 2006). Although post-apartheid policies have removed formal race barriers from schooling and introduced progressive language policies, the reality is that students from formerly African and Coloured schools are likely to be less well-prepared to enter university than their counterparts in former White and Indian schools. In the year of this study, in the BP and BHP courses less than half (42%) of the students were White, approximately 30 per cent spoke English as an additional language and 15 per cent had matriculated from historically disadvantaged schools. Scott, Yeld, McMillan and Hall (2005) have argued that UCT has to cater to a greater range of student preparedness than any comparable university in the world and, as the data reveals, this poses significant challenges for markers of student essays.

## **PROBLEM**

Within BP and BHP, essay assignments account for most of the in-course performance based and summative assessments. The essay assignments are judged against explicit criteria which are given to the students along with the assignment questions and to the markers together with content-based model answers. Since the inception of the course the practice has been for facilitators to mark the essays of the students in their groups. While there were certain advantages in this practice, over the years facilitators expressed concern about the potential 'biases' they brought to the marking of their own students. In order to strengthen the reliability of the marking, in the year prior to the study the course convener decided, with the support of the staff, that facilitators should not mark their own students' essays. In addition, a group of markers was selected from the BP/BHP facilitator team and given responsibility for all the marking. In implementing the new marking policy, a sample of essays was double-marked, once by the student's facilitator and once by a non-facilitator marker. The results of the analysis found no significant difference between the allocated marks. Despite these findings, it was felt that a more qualitative investigation of

markers' judgments was necessary in order to better understand the complexity of their decision-making processes.

The recognition of this complexity has been noted in a number of studies. Wyatt-Smith and Castleton (2005) found that 'There is no simple linear course that teachers follow to arrive at their judgements', and that teachers used 'official' and 'other (personal) factors' in reaching their judgements (135). Within the South African context, Reed et al. (2003) investigated intermarker reliability in the assessment of Honours research reports. They concluded that markers' interpretations depend on professional judgement, intellectual position and personal taste. Shay's (2003) study of the assessment of final year projects in an Engineering and Humanities department concluded there are a number of factors that influence markers, including their disciplinary orientations, years of experience and levels of involvement with students. This article is a further exploration of influences on markers' professional judgments, with a specific focus on judgments about language.

## **THEORETICAL AND METHODOLOGICAL APPROACH**

A number of assessment scholars have drawn on the hermeneutical tradition of Hans-Georg Gadamer to better understand assessment as an interpretive act (Moss 1994, 1998; Graaff, Banning and Shay 2004; Broad 2003). Moss (1998) uses the notion of the 'hermeneutical circle' to argue for an integrative and dialogic approach to assessment. Moss (1998) characterizes the 'circle' as representing a dual dialectic: one between parts of the text and the whole, and one between the text and the reader's foreknowledge, preconceptions or 'enabling' prejudices. Drawing on Gadamer, she makes a crucial distinction between 'blind' prejudices and 'enabling' prejudices (Moss, Schutz and Collins 1998, 142). The interpretive task is not to remove preconceptions – this is not possible – but to 'test them critically in the course of inquiry' (Bernstein 1983 cited in Moss et al. 1998, 141). Moss (1994, 7) argues, 'A hermeneutic (or interpretive) approach to assessment would involve holistic, integrative interpretations of collected performances that seek to understand the whole in light of its parts, that privilege readers who are most knowledgeable about the context in which the assessment occurs, and that ground those interpretations not only in the textual and contextual evidence available, but also in a rational debate among the community of interpreters'. This approach to assessment serves to explain how differences in interpretation arise as assessors privilege different forms of evidence and contextual considerations (Graaff, Reed and Shay 2004).

In order to explore markers' interpretive processes, data was collected by the lead author through two primary methods: interviews with the markers and observation of a marking moderation session. Semi-structured individual interviews were conducted with ten of the twelve BP/BHP markers who convened groups in 2004 and 2005. The interviews enabled the researcher to probe the markers' experience of assessing both their own students and students from other groups. The interviews were transcribed and coded for key themes. Markers were also observed and video-

recorded in a moderation session where markers are required to justify their marks and comments to other markers and come to consensus. In the analysis this moderation data was drawn on to both support and challenge the key themes which emerged from the interview data. These findings provide one 'window' into the complexity of assessment as social practice.

For the purposes of analyzing the data, the notion of the 'hermeneutical circle' served as a useful heuristic through which to understand some of the complexity of the markers' judgment-making process. The influences on markers is explored through the interpretive frame of four concentric circles (see figure) – with textual and contextual factors being explicitly delineated (Broad 2003). The inner circle is the 'object' of interpretation which includes both the student and the text produced by the student. The second circle or interpretive layer is the marker and all the 'prejudices' which they bring to the interpretive task. The third circle is the interpretive community of practice and the values which constitute this community. The outer circle is the institution and the competing discourses of transformation. It is important to note that these four 'circles' are distinguished for analytical purposes; in reality these influences on the interpretive process are deeply intertwined and inseparable.



Figure 1: Influences on markers

## FINDINGS AND ANALYSIS

### The text

As indicated in Figure 1, the innermost circle includes the textual factors that may potentially influence markers. 'Text' refers to both the written product and the BP/BHP student. The data revealed a number of aspects related to the quality of the essay that influenced assessors in their marking including evidence of understanding, clarity of knowledge, accuracy of referencing, ability to follow instructions, appropriate style of presentation and the most frequently noted area of difficulty for markers – that of language.

There was agreement among the markers that a well-structured essay using clear and correct English was an important criterion, but just how important and how much value to attach to it was debatable. What became clear was that on the basis of language, markers made inferences about whether a student was a first or additional language English speaker. In South Africa, issues of language and race are inseparable. This raises particular tensions for markers as prejudgments about language and race complicate the judgment-making process. One marker in an interview gave insight into these intra-marker tensions:

I gave this person 50, but the English was very bad. He had the basic understanding, but just couldn't get it across ... I don't like to hold that against people, because ... I think it was a Black student or someone who was writing in a second language, who has not had the same benefit of the education system that I have had. But yet understands the stuff, but can't actually express it. But it is very difficult because there is that side of you, and then there is the other side that thinks that well we are all supposed to be professionals and (students) are supposed to become professionals. And then their ability to write will be looked at and criticised by people who won't bear in mind those sorts of extenuating circumstances. I mean it is a very difficult tension there. So in that way my sort of political feelings and my general feelings about SA and its past influence how I mark, I think.

This marker gives insight into the inferential chain of prejudgments: from 'the English was very bad', the marker infers 'I think it was a black student ... writing in a second language', without 'the same benefit ... I had'. In trying to decide whether the student should pass or not, the marker articulates these tensions with honesty and clarity. In order to better understand these tensions they must be understood as an iterative movement between the first and second 'circles', the features of the text and the marker as interpreter.

### The marker

The second 'circle' points to the marker as interpreter. Gadamer (2003, 267–291) writes, 'a person trying to understand a text is always projecting' and 'all understanding inevitably involves some prejudice'. The problem is not prejudgments, which are inevitable; the problem is 'unfounded judgments'. According to Moss (1994, 7), this

hermeneutical perspective 'recognises that the reader's preconceptions, enabling prejudices, or foreknowledge, are inevitable and valuable in interpreting a text. They are in fact what make understanding possible'. Thus different preconceptions based on different values are inevitable.

The study of the BP/BHP markers revealed a range of prejudgements or prejudices. The BP/BHP markers are from diverse educational and professional backgrounds. Three of the markers, second language speakers, described themselves as being 'poor at English'. They admitted to assigning less importance to structure and language than other markers. They expressed empathy for those students who struggled with language. As one marker said:

I still struggle with the whole thing of language. How much should the student's ability to write English properly affect the mark ... Because I'm a second language speaker I have sympathy for that and empathy. You can see what he or she is trying to say doesn't exactly come out that way, and then you know it's quite difficult to allocate a mark to a paper like that. Because how much do you penalise a person for lack of good writing skills or English language skills?

The marker's question, 'how much do you penalize ...?' is a valid one. Crucial to the judgment making process is a wider interpretive community where these intra-marker tensions can be discussed and debated against a common set of values.

### **The community of practice**

A consistent theme emerging from the data is the importance of the marking team. The third circle points to this interpretive community and their collective interpretive process. In order to elaborate on this contextual influence, the insights of situated learning and in particular the work of Etienne Wenger (1998) are useful. According to Wenger (1998), a person's identity is made up of an interplay between individual and community aspects. This interdependent relationship resonates with the hermeneutical circle and the interplay between parts and whole. Wenger's (1998) concept of the community of practice sheds light on this interplay between the individual and community.

According to Wenger (1998) communities of practice exist because of mutual engagement, joint enterprise and shared repertoire. In terms of mutual engagement, individuals from very diverse backgrounds come together and organise themselves around a common purpose. Diversity is inevitable, but it is the relationships that are created between people that, despite differences, ensure the community is sustained. It is important that individuals feel and experience inclusivity. Complementary and overlapping roles must be acknowledged; tensions and conflicts must be managed through a process of 'community maintenance' (Wenger 1998, 74) where underlying dynamics are given recognition.

The BP/BHP markers identified the moderation process as this kind of community, a place where meaning could be negotiated. In interviews, a number



of markers spoke about the responsibility of taking the marking seriously and being 'committed to the assessment process' as well as having a 'loyalty to the team'. This resonates with what Wenger (1998) calls mutual engagement through supportive interpersonal relationships and the activity of community maintenance. This mutual engagement happens through joint enterprise, the process of negotiating meaning for the community and developing mutual accountability. For the BP/BHP markers the checks and balances of the moderation process make up this joint enterprise. For markers, the value of moderation seemed to be linked to both personal support as markers their difficulties and learned from each other's' experiences, and professional support as the moderation session provided a chance for the whole team to take responsibility for the marks. They valued the joint nature of the decision-making process. One of the markers commented:

The moderation process ... is the best thing. I find that in terms of the students' marks, we do really discuss them and people feel comfortable enough to say, "Hey, I don't agree with this". And that makes a huge difference. So I find it's very, very beneficial when discrepancies are picked up and then sorted out.

Finally, community membership and coherence facilitate the development of a shared repertoire. For example, a community of professionals develops a language or discourse particular and essential to their functioning as a community of practice. In BP/BHP the marking criteria formed the central discourse of practice. Markers had the criteria in front of them while marking, giving feedback and moderating. The criteria formed the basis of justification at each point.

Wenger's (1998) community of practice resonates with Gadamer's (2003, 302, 306) concept of 'horizon' which 'includes everything that an individual can see from their particular vantage point'. For any given individual there are 'limits on the possibility of vision'. Through the 'fusion of horizons' our own horizon is 'enlarged and enriched' (Bernstein, 183, 143). Through language and shared contact with the horizons of others, communities are forced to confront personal prejudices and resolve differences.

As previously noted markers struggled particularly with issues of language and how much significance to attach to the use of 'correct' English. The following data from a moderation session illustrates tensions as two markers struggled to agree on a mark:

- Janet: This is definitely a student who has a problem with English as a first language. She is from a rural area in the Eastern Cape.  
 I gave her 54 for that section  
 Lea: No, I gave her 70%.  
 Janet: That's because you know she has a problem and is a second language  
 Lea: English student. ... There was no true understanding. This could be because she doesn't understand the question.
- Janet: ... I think it's there, but she doesn't know how to get it out.

The two markers went on to discuss the student's work in great detail, comparing the information in the essay with the marking criteria. Finally they agreed on a mark of 58 per cent, but decided to hand the essay to other moderating colleagues for further scrutiny. This example indicates the importance of the community of practice as a source of support where individual prejudices are tested alongside the views of others who share a common enterprise and repertoire.

The example also brings into focus the underlying tension that exists for markers in terms of their role as group facilitator versus that of essay marker. Graaff et al. (2004) draw on the work of Paul Ricoeur to explore different hermeneutical stances from which assessors' judgment can be made. Ricoeur (in Graaff et al. 2004) distinguishes between two opposing hermeneutical stances in the social sciences – that of suspicion and faith. The stance of suspicion is a critical, judgemental perspective. The stance of faith is an empathic, sensitive approach. Both approaches recognise assessment as a form of judgment, but they are judgments from different stances.

These different judgment stances can be found in BP/BHP. The student-facilitator relationship, the course content and methodology could be perceived as a perspective of faith: the learning environment is student-centered and supportive, and a significant level of trust is encouraged. Facilitators are specifically chosen because of their interest in small group, experiential education with young adults. Course content is mentally and emotionally challenging since students are required to develop as knowledgeable, empathic and reflective health professionals. The focus in facilitator training is not only on the content for group sessions but also on understanding group process. Facilitators are encouraged to be empathic and sensitive to the students' learning processes.

In contrast, the BP/BHP assessment process could be perceived from a perspective of suspicion. Markers are required to step back from their empathic and supportive facilitator roles to become critical and judgemental. Explicit, pre-determined, objective criteria become the measures for performance. The written text is removed from its context of the group activities and group process. What was valued previously in group is back-grounded in the marking process.

Some of the challenges in the BP/BHP marking process may have to do with a struggle between these basic perspectives of faith and suspicion. Markers are required to move from a perspective of support and non-judgmentalism in their small groups to that of critical judgement as markers. In the previous example, Janet brought prior knowledge about the student to the marking process, notably that the student was not an English first language speaker and came from a historically disadvantaged schooling background. She struggled to separate this knowledge from the product in front of her. For Lea, her role as marker appeared to be one of judge as she brought Janet back to the text which from her perspective showed 'no true understanding'. While Lea recognised that Janet had information about the student, she did not allow this to influence the decision-making process. These inter-marker tensions are inevitable since different kinds of evidence are privileged from the perspective of different 'horizons'. They also are suggestive of competing value systems which operate at more macro-institutional levels, the final circle.

## The institution

The final outer interpretive layer of context represents the institutional, national and global imperatives for change in higher education. Given its history, the competing discourses of transformation are acutely felt in the South African higher education system. On the one hand, the transformation agenda at national level prizes issues of equity, access and re-dress of both institutions and individuals. On the other hand, to survive in a globally competitive system, the global agenda prizes quality, efficiency and excellence. This productive tension is noted in University policy documents which affirm the institution's aim to 'be a player in the global field whilst playing an active developmental and supportive role in its local African environment ... to ensure meaningful access and success for students and staff from diverse backgrounds ... and create an institutional culture where systems, processes, behaviour, symbols and rituals represent a diversity of culture' (Towards a Language Plan for the University of Cape Town 2005–2010, 2). The markers' deliberations suggest traces of these ideological tensions between the dual challenges of equity and excellence, tensions which are best resolved within a community where there is rational value-based dialogue about what really counts.

## DISCUSSION

Re-capping the main findings of the study, Gadamer's hermeneutical circle served as a useful heuristic through which to understand some of the complexity of markers' judgment making process. Influences on markers through the interpretive frame of four concentric circles have been explored. At the heart of the interpretive process, the inner circle, is the 'object' of interpretation which includes both the essay and the student. The data illustrated how even when the student is not personally known, markers infer identity from the text. Even the most explicit set of marking criteria cannot prevent these interpretations. The second circle is the marker and all the 'prejudices' which they bring to the interpretive task. Here the data illustrates how markers' own language backgrounds, their access to educational privilege and their 'political feelings' constitute (in part) the interpretive frame they bring to bear on their marking. This exposes intra-marker tensions which are never fully resolved even with the new policy of not marking their own students. The third circle is the interpretive community of practice and all the values which constitute this community. The data exposes the moderation dialogue as a site where competing discourses are both expressed and, to some extent, resolved as markers' individual prejudices are tested within the arena of the course and its values. The outer circle is the institution and the competing discourses of transformation. Once again traces of bigger Discourses (Gee 1996) are revealed in the intra and inter-marker tensions. It was noted that this analysis constitutes one 'window' on a complex social practice, and that the diagrammatic representation of such a complex process is inevitably limited. Nonetheless the findings are consistent with and contribute to recent

scholarship which emphasizes assessment as socially-situated (Gipps 1999; Filer 2000; Broadfoot 1996).

The most significant finding of the study is the reminder and reaffirmation of the role of the interpretive community in its task of making explicit the values which underpin its joint educational enterprise. As noted earlier, validity is no longer understood as a fixed property of assessment task or set of results, but rather the extent to which interpretations are sound for the purposes for which they are being used (Messick 1989). So on what basis can the BP/BHP community of practice be confident of the soundness of their interpretations? The findings of this article suggest that confidence cannot rest solely on evidence of reliability, as important as this is. Nor can it rest only on the standard practices of criteria, double-marking and moderation workshops. These are all important procedures, but they will not in themselves ensure validity. The basis for confidence will lie in the quality and the strength of what Moss refers to as the 'rational debate among the community of interpreters' (Moss 1994, 7). This value-based dialogue must be rigorous and robust enough to distinguish between 'blind' prejudices and 'enabling prejudices' (Moss 1994, 9), between criteria which are non-negotiable and those seemingly competing criteria which must be held simultaneously in tension. As Badat (2008) has argued, this is the challenge of transformation.

In the case of BP/BHP, a number of very practical strategies have been employed to strengthen the interpretive community and thus the validity of the essay assignments or validation process. With each essay assignment, clear marking criteria are made available to markers and students alike. While these criteria can never fully capture the full range of values, with each application the values of the course are put to the test. Written feedback is deliberately designed with links to the criteria so that students can verify strengths and weaknesses. Model answers are made available after the essays have been marked, and students are welcome to raise concerns through scheduled meetings with members of the marking team. In this way students and staff become socialized into a common though dynamically shifting set of values. The choice of markers is a crucial component of the validation process because they are also facilitators who bring knowledge, experience and insight into the marking experience. Markers are trained, supported and guided on an on-going basis that emphasizes consistency and accountability. Moderation is critically important in the validation process; it is here where the diversity of perspective is recognized and explored and where final consensus is reached – an illustration of Gadamer's (2003) fusion of horizons. In this way validation is an on-going community process that privileges dialogue about what is really valued.

## **CONCLUSION**

The findings of the study illumine some of the complexity of the judgment-making process; the multiple influences on the interpretive process. The acknowledgment that the assessment of complex performances is a socially-situated interpretive act means that

'assessment practices can only be understood within the social, cultural, and economic contexts in which they occur ... and the experiences, beliefs and expectations of both the assessed and the assessor constitute the meanings of assessment outcomes' (Shay 2003, 1). This has profound implications for the validity of assessment, about what is valued. The findings of this study serve as a sobering reminder of the responsibility of academic communities for robust and rigorous debate about what it values in student performances. Perhaps one of the greatest shortcomings of higher education is the absence of such debate. There simply are no shortcuts.

## ENDNOTE

The lead author of this article is the course convener for BP and BHP and thus has a key role in the development and maintenance of this community of practice. The research for this article was conducted by the lead author for the purposes of a Masters thesis in Higher Education Studies.

## REFERENCES

- Badat, S. 2008. The trajectory, dynamics, determinants and nature of institutional change in post-1994 South African higher education. Keynote paper presented at Higher Education Close up 4 Conference, 26–28 June 2008, Cape Town, South Africa.
- Bernstein, R. J. 1983. *Beyond objectivism and relativism: Science, hermeneutics, and praxis*. Philadelphia: University of Pennsylvania Press.
- Broad, B. 2003. *What we really value beyond rubrics in teaching and assessing writing*. Logan: Utah State University Press.
- Broadfoot, P. 1996. *Education, assessment and society*. Bristol: Open University Press.
- Cronbach, L. 1989. Construct validation after thirty years. In *Intelligence: measurement, theory and public policy: Proceedings of a symposium in honor of Lloyd Humphreys*. Urbana: University of Illinois Press.
- Filer, P. (Ed.). 2000. *Assessment: Social practice and social product*. London: Routledge/Falmer.
- Gadamer, H. G. 2003. *Truth and method*. Second edition (revised). New York: Continuum.
- Gee, J. 1996. *Social linguistics and literacies: Ideology in discourses*. Second edition. London: Taylor and Francis.
- Gipps, C. 1999. Socio-cultural aspects of assessment. *Review of Research in Education* 24:355–392.
- Graaff, J., Y. Reed and S. Shay. 2004. Validating academic assessment: A hermeneutical perspective. *Journal of Education* 33:51–68.
- Kapp, R. 2006. Discourses of English and literacy in a Western Cape township school. In *Academic literacy and the languages of change*, eds. L. Thesen and E. van Pletzen. London: Continuum.
- Messick, S. 1989. Meaning and values in text validation: The science and ethics of Assessment. *Educational Researcher* 18(2): 5–11.
- Moss, P. A. 1994. Can there be validity without reliability? *American Educational Research Association* 23(2): 5–12.

- Moss, P. A. 1998. Rethinking validity for the assessment of teaching. In *With portfolio in hand: Validating the new teacher professionalism*, ed. N. Lyons, 202–19. New York: Teachers' College Press.
- Moss, P. A., A. Schutz and K. Collins. 1998. An integrative approach to portfolio evaluation for teacher licensure. *Journal of Personnel Evaluation in Education* 12(2): 139–161.
- Nitko, A. J. 2001. *Educational assessment of students*. Third edition. New Jersey: Prentice Hall.
- Olckers, L., T. Gibbs, P. Mayers, M. Alperstein and M. Duncan. 2006. Early involvement in a multiprofessional course: An integrated approach to the development of personal and interpersonal skills. *Education for Primary Care* 17:1–9.
- Reed, Y., S. Granville, H. Janks, P. Makoe, P. Stein, S. W. van Zyl and M. Samuel. 2003. (Un) reliable assessment: A case study. *Perspectives in Education* 21(1), 15–28.
- Scott, I., N. Yeld, J. McMillan and M. Hall. 2005. Equity and excellence in higher education: the case of University of Cape Town. In *Equity and excellence in American higher education*, eds. W. Bowen, M. Kurzweil and E. Tobin. Charlottesville: University of Virginia Press.
- Shay, S. B. 2003. *The assessment of undergraduate final year projects: A study of academic professional judgment*. Thesis for the Degree of Doctor of Philosophy. Department of Education. University of Cape Town.
- . 2004. The assessment of complex performance: A socially situated interpretive act. *Harvard Educational Review* 74(3): 307–329.
- . 2005. The assessment of complex tasks: a double reading. *Studies in Higher Education* 30(6): 663–679.
- Towards a Language Plan for the University of Cape Town: 2005–2010. Discussion document. Language Development Group, Centre for Higher Education Development, UCT.
- Wenger, E. 1998. *Communities of practice learning, meaning, and identity*. Cambridge: Cambridge University Press.
- Wyatt-Smith, C. and G. Castleton. 2005. Examining how teachers judge student writing: an Australian case study. *Journal of Curriculum Studies* 37(2): 131–154.