

Carola Steinberg

Tom Waspe

DOI: <http://dx.doi.org/10.18820/2519593X/pie.v34i1.5>

ISSN 0258-2236

eISSN 2519-593X

Perspectives in Education

2016 34(1): 53-67

© UV/UFS



The struggle for academic rigour in assessment education

Abstract

This article explores the meaning of academic rigour in relation to a fourth year assessment education course for pre-service teachers. We present the requirements for a course to be considered academically rigorous, describe the course we offered in the light of these criteria and then present the students' responses. Our findings indicate differing perspectives between lecturers and students on what it means to learn about assessment and to be academically rigorous. Whereas the lecturers were expecting engagement with assessment theory and practice from all students, many students 'tuned out' whenever the course did not engage them in practical examples related to their subject specialisation. Only exceptional students moved beyond compliance with course requirements. The struggle for academic rigour involves developing a better alignment between lecturer and student expectations. This has implications for more explicit explanation of course purposes as well as increased cooperation with subject specialisation methodology courses.

Keywords: *Academic rigour, assessment education, alignment, curriculum, pedagogy, assessment, student perspectives, lecturer-student interdependence*

1. Introduction

A 2010 study of the experiences of the University of the Witwatersrand's (Wits) first year B.Ed. students negotiating their academic learning arrived at the claim that only students who achieved marks above 70% were able to adapt to the "new semiotic domain" of the university, where, compared to school, there is "less external locus of control and where knowledge is regarded as more than performance" (Shalem *et al.*, 2013: 1093). The students who were in the mark range of 40-60% (i.e. just below or above failure) were caught in a struggle for "epistemological access" and remained unsure of "how to differentiate between the words they need to select to explain a specialised idea. Furthermore, they are unable to identify the web of concepts within which their thinking is nested and how to select the textual evidence needed to explain their views and the examples required to demonstrate their point of view" (*ibid*: 1093). The form and criteria of academic discourse were so foreign to them that it prevented students from accessing the content that is

offered by university courses. For them, at the time, writing in ways that could be considered as academically rigorous was still out of reach.

It so happened that when we taught and evaluated the 4th year assessment course in 2014, we were working with the same group of students. Looking at the final course marks, only 52 out of 371 (14%) students achieved a mark of 70% and above, while 264 (71%) were in the 50-69% range. That made our pass marks look acceptable enough but was it any indicator of the course having been, and enabled the students to become more, academically rigorous?

This article tells the story of an assessment education (AE) course in which the staff intended to offer epistemological access to a course that was academically rigorous, which, at the time, we understood to mean high quality teaching and learning. We saw it as our responsibility to offer clear, explanatory lectures on key assessment concepts and debates, appropriate readings, challenging and relevant tasks, useful feedback, and explicit guidance for student self-regulation of their work, so that students could acquire key assessment concepts and the accompanying professional skills. From our perspective, the students' responsibility was to engage with the lectures and ask questions, read the provided texts with understanding, put effort into writing the set tasks and generally enjoy the learning. This intention requires effort all round, which we conscientiously put in. Yet, to our dismay, we found that many students did minimum work and some evaluated the course as being 'boring', 'irrelevant' and 'useless'. Rather than losing faith in ourselves and our students, we decided to embark on the road of academic rigour by engaging in "collegial conversations that encourage deep and critical reflection for teacher educators" (Selkrig & Keamy, 2015: 1) The question that drove this conversation was: What are students' perspectives about what and how they learned from the assessment course?

2. A conceptual framework for understanding the term 'academic rigour'

The term academic rigour (AR) is used in the context of a general discourse about the quality of education. While there are many variations of the definition, no decisive consensus regarding a definition has yet emerged (Gray, 2008; Lincoln, 2010; Blackburn, 2013; Draeger *et al.*, 2013; Reich, Turner & Volkan, 2013). Despite this, there is not much *debate about* the definition; it tends to be defined and used to suit the context within which it is applied. Nevertheless, there are shared facets that characterise the discourse of AR.

AR concerns the three message systems of education: curriculum, pedagogy/instruction and assessment (Blackburn, 2013; Lincoln, 2010). "True rigour is weaving together the elements of curriculum, instruction, and assessment in a way that maximises the learning of each student" (Blackburn, 2013: 13) and to be achieved, the three message systems need to be aligned (Ainsworth, 2011). Alignment operates along two axes: breadth alignment ensures that all the relevant content is meaningfully covered, while depth alignment, which can be achieved by, for example, using Bloom's taxonomy (Gray, 2008), ensures that the curriculum and pedagogy reach pertinent levels of cognitive demand based on challenging learning objectives. Thus rigorous and aligned curricula, pedagogy and assessment ought to "promote in-depth learning and the use of cognitive skills similar to those found in the higher order thinking levels of Bloom's taxonomy" (Reich *et al.*, 2013: 6).

Academically rigorous and aligned education infuses the three message systems with high expectations. "Holding high expectations for student learning is at the heart of academic rigour" (*ibid*: 8). These expectations of student performance should be highly visible and be made "explicit in course syllabi, rubrics and assignment directions" (Gray, 2008: 5). Yet high expectations require a particular attitudinal approach when teaching: "Academic rigour ... is more likely to exist in schools with cultures that foster high expectations of all students and that have an overall focus on providing students with educational experiences that challenge them" (Reich *et al.*, 2013: 19). Achieving these expectations involves presenting meaningful, challenging and relevant content with high cognitive value that focusses on core concepts and the "big ideas" of the discipline (*ibid*: 10) as well as being aligned with the official curriculum as mandated by policy. It also involves interactive teaching and active learning focusing on higher order thinking (Lincoln, 2010; Blackburn, 2013; Draeger *et al.*, 2013). To achieve this, methods and activities should be appropriately selected from a repertoire containing multiple possibilities (DeLuca *et al.*, 2010) to suit the content. This combination of "content-based teaching" complemented by "process-based pedagogies" "seeks to engage students in active meaning-making through processes of critical reflection, dialogue, and experiential and authentic learning" (DeLuca *et al.*, 2013: 130). The third component, rigorous assessment, entails quality formative and summative assessment wherein assessment criteria are explicitly explained. The design and selection of aligned and appropriate assessment methods covering both breadth and depth, theory and practice ensures the validity, reliability and fairness of rigorous assessment practices (McMillan, 2014).

Arising out of this literature review, we are working with the following understanding: academic rigour entails high expectations of student attainment of deep learning which spans theory and practice. It is manifested through a curriculum which balances breadth and depth of knowledge (with an emphasis on depth) while being aligned with a supportive pedagogy, quality assessment and being appropriate to the knowledge of the discipline and the teaching context. Therefore, to demonstrate in what ways the assessment education course we offered was academically rigorous in the above way, we need to show how we adapted these criteria (student attainment, deep learning, theory and practice, breadth and depth of knowledge, supportive pedagogy, quality assessment, appropriate to discipline and context) to its particular constraints.

These criteria for AR are an ideal model and are crafted from the perspective of teaching and curriculum design. They also posit an, often implied, ideal role for students. This ideal role includes effective self-regulation, significant effort, commitment and motivation. Empirical studies (Draeger *et al.*, 2013) have indicated that teachers and lecturers would concur with the model presented above, however, studies with students paint a contrasting picture, with differing conceptions depicted (Draeger, del Prado Hill & Mahler, 2014; Mahler *et al.*, 2014). Lecturers were mostly concerned with how quality learning could be attained while students' overriding concern was how hard it would be to get good grades. Whereas lecturers emphasised active learning and participation, students were mainly concerned with workload. Lecturers stressed the importance of higher order thinking while students focused on the difficulty of the material. Lecturers had expectations of teaching meaningful content however, students worried about the instrumental utility of the content for their future careers. Lecturers stressed higher order and critical thinking but students were concerned about grades and the degree of difficulty in attaining high standards. It is these divergent representations of AR, which sets it up as an area of contestation and terrain of struggle. So the specific question

that arose for us was: What tensions arise when attempting to ensure academic rigour in an assessment education course?

3. Presenting an academically rigorous assessment education course

Assessment education (AE) is a relative newcomer to undergraduate teacher education programmes. Yet the need for a formal introduction to assessment theory and practice is supported by professional demands on teachers, by policy mandates and by research reviews. Stiggins (1999) pointed out that teachers spend an estimated 30-50% of their professional time engaged in assessment activities (DeLuca *et al.*, 2010: 20). In South Africa, national policy mandates assessment knowledge and skill as a professional requirement of teachers (Department of Basic Education, 2012a, b). Additionally, a research review study conducted in the USA found that “explicit assessment education at the pre-service level has the potential to support positive changes in teacher candidates’ conceptions of assessment and promote their assessment literacy” (DeLuca *et al.*, 2013: 129).

The question then becomes, how can AE for undergraduate students be made practically appropriate and academically rigorous? It could be argued that “given the constraints of pre-service teacher education programmes – comparatively short on-campus cycles (which does not apply in our case), competing demands on content and often large-group instruction (which does apply) – it is unlikely that candidates will complete their pre-service year with a robust and comprehensive understanding of assessment practice, theory and philosophy” (DeLuca *et al.*, 2010: 25). Yet if AR is understood as being relative to the purposes, objectives and constraints of a course, then the general criteria for AR can be adapted to fit a particular AE course.

3.1 The course context

In response to the policy mandate that teacher education should develop teacher competence in assessment knowledge and practice (Department of Higher Education and Training, 2011), the re-designed B.Ed. of the Wits School of Education includes a 6-week assessment module in its 4th year. As a module within the education theory major, all students are required to attend, which means the assessment course caters for student teachers from all phases and subject specialisations.

Prior to developing the course in 2013, we met with colleagues across the school and agreed that this course would provide key assessment concepts, principles and skills that remain appropriate regardless of educational level and subject matter, while methodology courses would elaborate on and adapt assessment principles in subject and phase specific ways. For the assessment course, it was decided to use examples from language and maths in the inter-sen phase, which we hoped would be general enough for all students to relate to. The intention of the course was to provide a concept-based approach to assessment, which would provide students with a principled way of thinking about the dilemmas posed by the practical concerns of assessment.

3.2 The course content

The first few lectures in the course fulfilled the requirement of presenting the knowledge and enabling understanding of the “core concepts of assessment theory” (Rudner & Schafer, 2002).

First came the educational purposes of assessment: the tension between the traditional purpose of using assessment to establish the amount of learning compared to the educational purpose of using assessment to generate more learning, i.e. assessment for and of learning and using assessment for accountability purposes (DeLuca *et al.*, 2010). Then came the imperatives underlying all quality assessment: the need for ensuring the reliability, validity and thus fairness of assessment, whether it was for formative, summative or accountability purposes. The second part of the course moved to the development of professional assessment practice (DeLuca *et al.*, 2013) and its “authentic problems” (Reich *et al.*, 2013: 10). It covered teacher responsibilities such as evaluating and generating varied, clear, reliable and valid assessment tasks with marking schedules/rubrics, doing error analysis, understanding assessment policy and the need for record keeping, collecting and using assessment data for school improvement and the thinking behind the system-wide accountability testing in the form of Annual National Assessments. At the same time students learned common assessment vocabulary/concepts, such as content/construct validity, criteria, taxonomies, levels of cognitive demand, marks, averages, means and standard deviations. The course ended by presenting research on teachers’ assessment emotions.

By focussing on concerns of educational purpose and professional practice, the course intends to engage students in a productive dialogue between theory and practice. The conceptual sections intend to distantiate students (Slonimsky & Shalem, 2006) from their taken for granted experience of having been assessed and also to offer them a lens through which to recognise the value (or not, depending on purpose) of the more technical skills of assessment they are about to learn. The practical sections intend to give students an insight into the complexity of considerations involved in generating assessment tasks and judgements, so they develop “positive professional dispositions” (Stiggins, 1999) towards assessment and have a sufficient foundation to learn from school conversations as new teachers.

By limiting the key foci of the course to two core concepts (i.e. educational purpose and assessment quality) that re-emerge throughout the professional responsibilities, we hope the course generates enough breadth, yet with an emphasis on depth.

3.3 The pedagogical means

(DeLuca *et al.*, 2013) suggest that AE should use “multiple strategies” which include “content-based teaching” and “process-based pedagogies” (*ibid*: 130). That description fits our thinking precisely. Our intention was to use content and pedagogical means to inform and inspire students and regulate their learning.

Content-based teaching refers to “didactic instruction with a focus on the transmission and application of knowledge and skills, typically through lecture-based, text-based or case-based learning” (*ibid*: 130). Over 6 weeks we conducted 10 double-period large class lectures with some interaction enabled by the use of ‘clickers’¹, plus 2 computer lab sessions, during which the lecturers were present to help students. Each lecture was accompanied by 1-4 academic readings (both local and international). Students received a detailed course outline with tasks and academic readings, on paper and through a course website, which also contained all the PowerPoint presentations. They had office numbers and email addresses of the three staff members and came for consultations when they had administrative or academic questions. However, due to insufficient staff in relation to students, in 2014 it was not possible to offer

1 Clickers are devices used in classroom response systems

tutorials for subject specialisation groups of 30-40 students, as we had during the first year of the course.

We complemented the lectures with “process-based pedagogies” that “seek to engage students in active meaning-making through processes of critical reflection, dialogue, and experiential and authentic learning where the art of assessment would be practiced” (DeLuca *et al.*, 2013: 130). Our students needed to complete 5 out of 7 thoroughly explained and “analytically scaffolded” (*ibid.*: 133) weekly tutorial tasks for online submission, which, between them, also spanned theory and practice. To make students more “aware of their own thinking on assessment” (*ibid.*: 134) they were encouraged to work in pairs. The first task was conceptual, asking students to describe formative assessment and served as a preparation for the essay assignment. The remaining tasks used a “strongly demarcated practical context” (Shalem & Rusznyak, 2013: 1125) to give students a practise run of enacting teachers’ professional responsibilities. Students were given the opportunity of preparing an assessment task and uploading it onto test design software, comparing different versions of assessment criteria, recording and reporting marks, reflecting on and discussing marks with a teacher during teaching experience, analysing ANA statistics and doing error analysis. Although these tasks had the potential to be (and in 2013 functioned as) “multiple-perspective conversations” that allow for engagement “in conversations with peers about readings, assessment scenarios, and dilemmas of practice” (DeLuca *et al.*, 2013: 133), we think that in 2014 they were less effective. This is because with students working on their own or in pairs, the lecturers’ perspectives were missing from the discussions, making students more concerned with meeting the submission date than with discussing dilemmas raised by the tasks.

3.4 The assessment format

Central to rigorous assessment in AE is the modelling of “sound assessment practices as a matter of routine in their own course assessments” (Stiggins, 1999: 24). Formative assessment should be modelled while at the same time “explicitly instructing on assessment concepts and practice”, a combination which empirically has been found to be “highly supportive for student learning” (DeLuca *et al.*, 2013: 137). DeLuca (2013) terms this “modelling through assessment pedagogy”, where diagnostic assessment as well as ongoing giving and receiving of feedback are fully integrated into instruction as assessment for learning. The formative assessment of our course was embedded in the tutorial tasks described above, as the tasks needed to be submitted online at the rate of one a week, so as to receive the ‘due performance’ necessary for permission to write the exam. We recorded all the submissions but were unable to give regular feedback.

DeLuca suggests this modelling be carried through into summative assessment, where a “congruent approach between concepts taught about assessment and the mechanism for grading teacher candidates’ work” (DeLuca *et al.*, 2013: 137) is practised. Assessment practices in AE programmes must follow the quality assessment criteria of validity, reliability and fairness (McMillan, 2014). The summative assessment of our course consisted of an essay assignment and a sit-down examination – both worth 50% of the course mark. The essay asked students to engage with the purposes and methods of summative and formative assessment using a range of readings. The exam offered a choice between two essay topics – error analysis and task design – and used various short question formats to cover the course curriculum.

4. Course evaluation: Noticing the gap in expectations with regard to AR

It is common place for student evaluations to be used as a measure of AR including lecturer effectiveness (Clayson, 2005; Stark & Freishtat, 2014). Accordingly, and as per institutional requirement, we conducted a student evaluation of the course. Items for the evaluation were selected from a list provided by the Wits Centre for Learning and Teaching Development (CLTD), who processed the survey and provided results. The selected items related to key aspects of AR, namely teaching, learning, content, relevance, coherence, student interest, alignment, workload, effort, grades and expectations.

The results of the survey were disappointing. CLTD informed us that the course was rated 1.1 points (out of 10) below the institutional average and one qualitative comment said “it was one of the worst education courses ever presented!” Although there were several positive comments to balance the negative, this comment galvanised us into further investigative action.

We soon realised that the survey results were problematic for a number of reasons. Firstly, we conducted the evaluation after the final grades were released to the students. This was a mistake because “students punish instructors for low grades” (Clayson, 2005: 10) and “people tend to be motivated to act (e.g. fill out an online evaluation) more by anger than by satisfaction” (Stark & Freishtat, 2014: 2). Secondly, the response rate of the evaluation was low, only 22% (80 out of 371). “The lower the response rate, the less representative the responses... and there is no justification for assuming that non-responders are just like responders” (*ibid*: 2). This means that “if the response rate is low, the data should not be considered representative of the class as a whole” (*ibid*: 3). Thirdly, institutional “averages and comparisons make no sense, as a matter of statistics” (*ibid*: 3) because it is unclear what underlying constructs are being measured. Fourthly, student evaluations “do not measure teaching effectiveness. We measure what students say, and pretend it’s the same thing” (*ibid*: 5). Even “reliability has little to do with whether evaluations measure effectiveness” (*ibid*: 6). Instead, it becomes a question of validity, are the student evaluations measuring course rigour or are they measuring some other construct such as satisfaction with course grades?

This goes back to the disparity and gap between lecturer and student expectations discussed earlier. Students expect high grades with minimum effort and lecturers expect learning with high effort. “Students have high expectations for rewards” (Clayson, 2005: 4) and rigour “is not within their perceptual norms” (*ibid*: 10), indicating that “students did not believe that a demand for rigour was an important characteristic of a good teacher” (*ibid*: 11). Thus, how should rigour be evaluated and measured? Learning is often used as a proxy for rigorous teaching but measuring learning is hard (Stark & Freishtat, 2014). Course grades are often used as a measure for learning and AR but it is “not clear that higher course grades necessarily reflect more learning” (Stark & Freishtat, 2014: 5). “In general, to infer causes such as whether good teaching results in good evaluation scores requires a controlled, randomised experiment” (*ibid*: 6).

For these reasons we are not reporting on the averages from the survey as measures of AR. “Instead of reporting averages, we should report the distribution of scores for instructors and for courses: the percentage of ratings that fall in each category” (*ibid*: 4).

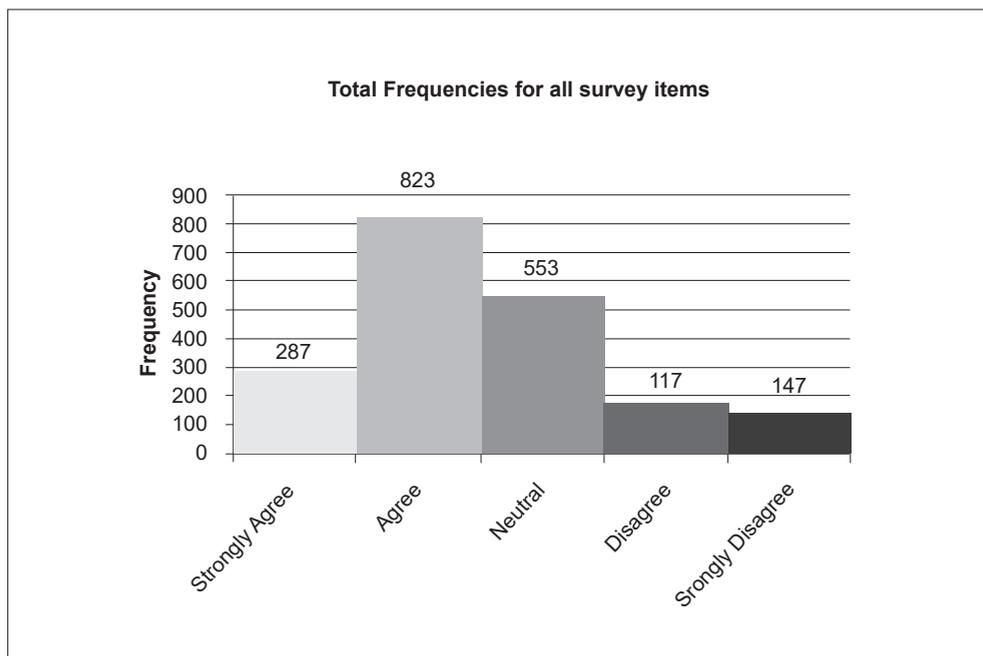


Figure 1: Total frequencies for all items

When we aggregate the 'strongly agrees' and the 'agrees' we get 1110 (56%) compared to 324 (16%) when the 'strongly disagrees' and 'disagrees' are aggregated, with 553 (28%) neutrals. This provides more specific information than the comparison against the institutional average. Yet we still cannot infer that the course was academically rigorous. At best, it may reflect a certain degree of overall satisfaction with the course as expressed by the students.

5. Digging deeper: Students' perspectives on the AR of the course

Following due ethical procedures, we sent out invitations to participate in focus group discussions to all students. Eighteen students, nine young men and nine young women, participated in one of five focus group interviews, which lasted approximately an hour each, were recorded and transcribed. It was a self-selected sample, whether students were positive or negative about the course, they came because they were more concerned and engaged.

5.1 Students' responses to the curriculum

When asked what they had learned from the course, students in the focus groups² responded with a range of issues and between them, they covered all topics and tasks in the course. When asked what they disliked about the course, much the same thing happened. This means we received useful feedback for course improvement from students but no recommendation for removing topics from the course.

2 We obviously cannot say what all students learned. A self-selected sample of 18 out of 371 students can provide insight but not generalisations.

The reasons students gave for their choices arose primarily from their practical focus on becoming a professional. They paid attention to the course because “with assessment, it’s what you do every day as a teacher” and they approved of the course because it provided “very, very useful information” that was “more practical” than other education courses. They liked that the course went “hand in hand with the subject methodology” courses because “then at least I know, this is making me into a teacher”. They liked designing assessment tasks electronically because “it would be useful in terms of saving time and paperwork”. They liked practising skills that “could come in very handy next year” and doing tasks that “you can actually go to schools and implement this”. This practical focus generated “excitement” and was a motivator for engagement with the course and the tutorial tasks.

Some of their insights into assessment issues spanned theory and practice and enabled students to make a shift in perspective, to “look at it (be ‘it’ task design, error analysis, or ANAs) from a different view”. For example, students expressed their approval of learning about error analysis not only because it was ‘new’ and ‘interesting’ but also because it made them realise that “you have to give the learners feedback”. This was something they had “not taken seriously” before and because they learned that “maybe the things you say as a teacher might disadvantage the learners because they are a misconception of the teacher”. They commented on having gained a more comprehensive understanding of the ANAs whereas before, the ANAs had been viewed as “ah, these people (the education department) are just giving us a lot of work”, students could now see “how the ANAs are also doing research” which helped them to “agree with the stats that are in place”. The most dramatic description of a shift came from a student who had complained about the assigned essay comparing summative and formative assessment and who received a revised topic that required him to make an argument about why summative and accountability testing persists in the face of evidence that it does not promote learning. This “opportunity to delve deeper” pushed him into “actually looking at what we were studying”, which proved to him that he “was able to do that sort of thing” and filled him “with such a deeper knowledge of what summative assessment and formative assessment is and how it affects our classrooms”, which meant that the later “lectures started making a lot more sense”. Their practical interests in the processes of assessment made it possible for students to gain an understanding of the educational purposes and conceptual complexities that assessment presents to teachers.

Yet the practical focus could also be an inhibitor to learning. Expecting the course to be practical lead to rapid demotivation whenever students “didn’t see the whole point of doing it” and generated impatience with any theoretical exploration of purpose or methods that did not provide examples directly related to their subject or phase specialisation. The most frequent reason for students wondering, “why am I doing this?” and then “tuning out” and “being bored”, was when they could not see a direct relationship between assessment and their subject/phase specialisation.

Our plan of using examples taken from inter-sen maths and English for lectures and tutorial tasks did not work well enough. Students from FET, foundation phase, technology, history, arts and culture and science complained that they could not relate to the examples: “So the topics as such were useful, but the examples were not, because they were all maths and English, so you thought ‘this is a maths lecture, and I’m going to tune out’”. Students in the focus groups strongly advised that the course should “use different subject areas for the examples, not just focus on maths and English”. During 2013, we had placed students into tutorials according to

their subject and phase specialisation, which enabled discussion of specialisation concerns in relation to the maths and English examples provided. However, with the decrease in staff in 2014, the tutorial activities were turned into tasks for submission, which dramatically lowered the potential for explanatory discussions. Seen through the eyes of the students, it was a lost opportunity for them to engage with the specificities of their specialisations. Be it “using analysis skills in history tasks”, “how to look at and assess technology models”, “how to do error analysis in social sciences” or “understanding how the impact of the ANAs done in primary schools filters through to high school students”. One student spoke for all when she politely insisted that “I just feel if you’re teaching a course that, you know, everyone is doing, then it should sort of relate to everyone, even if just a little bit”. The relationship between assessment principles and specialisation specificities requires much more across-courses collaboration between subject methodologies and assessment in the future – so that assessment has a greater range of examples, the specialisations remind students of key assessment concepts, and the staff needs to know enough about one another’s courses to cross-refer or even set joint projects.

5.2 Students’ responses to the pedagogy

Students approved of the technological aspects of the pedagogy. They liked the online submissions because “you were moving with the times”. They liked the electronic tasks even though students who were less “digital” experienced it as “very tricky” and had to “do it over and over again”. They debated whether it was better to learn these skills “by themselves, through trial and error”, or whether the instructions should be written out “step by step by step”. Yet they agreed that “the electronic section of the course was well done” and they liked the “freedom to be left in the computer lab to do work without being monitored” (although staff were available for support). Regarding the clickers, they were unanimously positive. “Keep the clickers, the clickers were exciting, so much fun” was the consensus. And for good reason: “The lecturer asked questions out of the blue and that would get us all active and working and engaging with the lecture”. “We could see the response actually on the board there – that was thumbs up for me”. “We could interact in the class, feel part of the lecture and have a voice in saying “this is more important than that”. And, in terms of the assessment conundrums they had to respond to, “when the explanation was given, I never forgot the things I got wrong”. The clickers created a space where students’ knowledge and opinions were recorded and the complexities of assessment issues could be discussed.

The lectures were more ambiguously received. Some students found the information provided “brilliant”, “useful”, “insightful”, and “enjoyed” the lectures because they were “very interactive and close to what we do, not theoretical”. Other students found the lectures ‘boring’ and said it was too much “to expect 350 people to sit very quietly for two hours”. Students had contradictory reasons for their critique: for some, the lectures “were a little bit too explanatory and repetitive”, while others wanted “more detail”. Many would have preferred smaller tutorials after each lecture so as “to elaborate on some of the ideas and concepts that were covered in the lectures”. As one student argued:

I think we should have more contact with lecturers and with tutors, and have more... how can I say it...like a close-up encounter with the lecturers. Because we are doing a professional degree, we must learn from people who have done it before. So just to be lectured, the students get bored, and those who are not bored, get bored by those who are making noise.

The lecture venue worked against any “close-up encounter” – the battery in the roving microphone went flat quickly, the acoustics amplified student chatter, the rows of seats were long and required many students to get up if one student needed to move. Having five different lecturers, all experts in their fields, deliver the lectures, did not make an impression on the students either way. They blamed the lack of interest in the lectures on their being in the last year of their degree: “the problem is sometimes, I don’t know, like, when you get to fourth year, we kind of run out of steam or we get tired and we complain almost about everything”.

5.3 Students’ responses to reading

The course expected students to engage in regular reading prior to each lecture. In the open-ended questions on the evaluation forms, students most frequently complained that there was “too much reading” and they wanted “less reading”. In the focus groups, they elaborated on their reasons: “I hated the course when I found that book. This is the biggest book I’ve ever had in all my life”. “Just by looking” at the collection of articles they felt “scared”, “intimidated”, “too tired” and “lazy to read”. One student described how he “read the readings for 10 minutes before the lecture” when actually, he was talking about skimming the lecture PowerPoint. Several students admitted that after they had read the first 2 or 3 articles “it was really hard to keep up with the readings for the lectures” and there “were a lot of readings that I ended up not reading”. When asked whether they would rather find their own texts to read, the students responded negatively: “No, it’s better if you just give me what I need to do. Students work better when they’re given the exact thing they need to do. Looking for yourself what to read and do, no, that is a hideous task. It’s better to just work with what you have”.

Given this general reluctance to read for information, students tended to exercise choice. They skimmed through the book and chose articles they found “interesting” (e.g. “how do you design a multiple choice or a true/false question?”), “relevant”, “useful” and, above all, “practical” (e.g. “not case studies from Europe but ideas for doing things in Tembisa”). Some admitted that “the readings were not that bad”, “not as long as we thought they would be” and “there were nice ones that I learned a lot from”. One student even mentioned articles that presented an academic argument and how this might be a stimulus for reflective thinking:

The articles on summative and formative assessment were really helpful, particularly Black, as the main argument, because after reading Black, and the other guy... Shepard, you got two different views. Then you can be either neutral or critical.

When they did read, some students found their efforts were rewarded. As one student discovered, “when I finally sat down and did the readings, I felt they were insightful, easy for me to read and understand, and I actually understood more than I had understood in the lectures”.

The focus groups made it clear to us that the students were not uncooperative when they complained about “too much reading”, but rather that regular reading was not a part of their culture of learning. It required external pressure for them to make the effort to engage in the task. One student passionately made the point that “the Wits Education campus does not have a culture of working ... hard enough” and that “we’re not required to produce our best work in order to pass and so learn to work harder”. He wondered “whether that is the subject’s fault, the lecturer’s fault, the university’s fault? Where does the accountability come in for the actual people who are studying that degree?” He argued that it was not sufficient for particular courses to raise and consistently enforce the demands they placed on students but that “it’s

got to change across subjects” because “there’s something there that’s so deep, that’s got to kind of be flipped around”. Using the term “institutional consistency”, he pleaded that “this university needs to create some sort of culture of learning”.

Breadth and depth of content coverage and understanding is not possible without sustained student reading. Yet the students’ perspective expressed both resistance to and an understanding of the need for reading. It left us thinking that it is important to generate a conversation about the culture of learning and reading with colleagues across the school. What pedagogical means could build a culture more conducive to reading within and across courses?

5.4 Students’ responses to the assessment

The focus groups enabled students to reflect on how they functioned in a “culture of not pushing ourselves to work hard enough”: “I only do readings when I have a tutorial and I wait for somebody to tell me: okay, you must do your readings”. The prime motivator for learning was the assessment. “It’s like, why did I do this if it’s not going to be assessed?” “If it’s not for marks, we don’t take it seriously. That’s how we’ve been trained”. They responded to the assessment demands of the course by making strategic choices for learning: “As soon as I receive the coursework, I check if there is a question for the assignment. Often the question states clearly what you should read, so I focus on that alone. Actually, that’s all of us.” This made students dependent on the regulative mechanisms built into the course to push them into making an effort. “The Wits-e thing worked a lot, because it helped me to study, because I had to submit something. If we don’t get the task, as students we won’t study”. They abdicated their responsibility for learning not only to the course but also to the lecturers: “If I don’t get a comment, so why should I bother writing the next tutorial? I was waiting for the previous comment to come before I can proceed to the next tutorial”. At its worst, this sense of being reluctantly dragged into learning through the demands of assessment and led students into a compliance approach that side-lined the subject matter being learned: “I will do cram work so that I pass. You just want to cram, pass and forget. And when the student goes to the next grade, they know nothing. They will even say, no, we didn’t do that last year (laughs). Whatever we do, it’s not really here (in the head).” Which made us first wonder why we bother making the effort to create what we think are interesting and academically rigorous courses if nothing we teach remains in the students’ heads – but then we became even more determined to adjust the course so that students cannot but engage with it.

6. The tensions of co-generating a culture of learning: The interdependence of lecturers and students

If we understand AR as a two-way interaction, where the rigour of the lecturers needs to be mirrored by the rigour of the students, then it is not enough to understand how students respond to a course. It also becomes important to understand how students develop their ability to learn during the course or how the course fosters that development. Therefore, we analysed what students said about how they motivate and regulate their own learning.

As fourth year students who had spent many weeks teaching in schools over the years, this conversation about their lack of motivation for learning enabled a reflection on the difference between their motivation to learn when being in student mode compared to being in teacher mode.

I think we as students, we become learners when we're on campus and teachers when we're on teaching experience. When I'm on TE, I want to go and research for my preparation and my lessons, I want to get more information, come to the library, search for that. But when I'm on campus as a student, doing all of that is too much; I don't want to do it. I get into student mode. And when I'm on TE, I get into teacher mode. Yes, that's what happens with me.

This insight came when it was too late for students to change anything – it was the end of their fourth year. Even then, no student indicated that it might be a good idea to get out of “student mode”. They had been in that role for at least 16 years of their educational lives. The formative assessment described above generated a tension between compliance and learning for them, some students opted for compliance only (“just hand in whatever on that day”), while others used being “scared of the DP thing” as motivation to engage and “learn some things that we didn't know before”.

The way in which students simply assumed that their motivation for engaging in learning activities depended on course regulations and incentive provided by lecturer feedback, places a double educative responsibility on courses that intend to be academically rigorous. As a baseline, academically rigorous AE courses need to provide clear, comprehensive, accurate and context appropriate information about the topic at hand. In addition, they need to generate AR among their students. We would argue that this involves two aspects. The first aspect is to provide course structures that pressurise students into regularly doing academic tasks that enable learning, the second is to provide inputs that gradually wean students into an attitude of taking responsibility for their own learning. For students to develop agency and become more rigorous academically, courses require a well-explained structure, which needs to be complied with and enables choice.

7. Final reflections

It would appear that although learning happened and the survey, as well as the focus groups indicated a general sense of satisfaction, it is questionable whether the assessment education course achieved deep, rigorous learning on the part of the students (see the marks in section 1 above). The high expectations of the lecturers and the shallow learning of the students generated tensions in the pursuit of academic rigour.

The key principles of academic rigour, as identified in the literature, are high expectations, alignment between curriculum, pedagogy and assessment, balance between breadth and depth and for assessment education, an integration of theory and practice (Gray, 2008; Ainsworth, 2011; Reich *et al.*, 2013). In the assessment education course under discussion, there were tensions that emerged regarding three of these four principles. No student mentioned problems caused by a lack of alignment between the curriculum, pedagogy and assessment components of the course. However, there were tensions regarding the other three principles.

Regarding high expectations, both the literature and the data pointed to the realisation that the issue is not a case of lecturers having high and students having low expectations – both groups have high expectations but of differing outcomes and processes. Lecturers tend to have high expectations for deep and broad student learning while students tend to have high expectations for a favourable trade-off between the lowest possible effort and attaining the desired mark (Draeger *et al.*, 2014; Mahler *et al.*, 2014).

Regarding the other two principles, the course intended to achieve a balance between breadth and depth, theory and practice. It covered a basic range of assessment concepts and practices (Stiggins, 1991; Rudner & Schafer, 2002; DeLuca *et al.*, 2010) both at a conceptual/theoretical level and by enabling depth through activities of application (DeLuca *et al.*, 2013). Yet, from the perspective of the students, the activities offered did not sufficiently take them into the assessment practices of their specific subject and phase specialisations, thus inhibiting their depth of understanding and engagement with practice. The course needs to engage with this design issue to better meet students' needs. Another tension that arose was the volume of reading required. Reading is essential for breadth and depth in theory and practice. Yet, students' tendency to limit reading to what is assessed did not match well with lecturers' expectations of reading for deep learning. In our opinion, this lack of engagement with text is the biggest barrier to academic rigour, which would require a school-wide discussion about fostering a culture of reading.

There is an unavoidable tension between lecturer and student expectations (Draeger *et al.*, 2014; Mahler *et al.*, 2014) which is constitutive of academic rigour. This research enabled us to come to an understanding of these tensions and left us with a new question: How best can these tensions be brought into a creative relationship with one another?

What is needed is more congruence between the expectations built into the course design and delivery on the one hand and the expectations of students and the role that they play on the other. The struggle for AR entails resolving the divergent expectations between lecturers and students. Finding our way through the struggle will involve, at a minimum, clearer explanations to be given to students regarding the assessment course purposes, collaboration between assessment staff and subject specialisation methodologists to offer a wider choice of tasks and a school-wide discussion about fostering a culture of reading.

References

- Ainsworth, L. 2011. *Rigorous curriculum design: How to create curricular units of study that align standards, instruction, and assessment*. Englewood, Colorado: Lead+Learn Press.
- Blackburn, B. 2013. *Rigor is NOT a four-letter word*, 2nd ed. New York: Routledge.
- Clayson, D.E. 2005. Academic rigor: A critical analysis. *Universitas*, 1(1).
- DeLuca, C., Klinger, D., Searle, M. & Shulha, L. 2010. Developing a curriculum for assessment education. *Assessment Matters*, 2, 20–42.
- DeLuca, C., Chavez, T., Bellara, A. & Cao, C. 2013. Pedagogies for preservice assessment education: Supporting teacher candidates' assessment literacy development. *The Teacher Educator*, 48(2), 128–142. <http://dx.doi.org/10.1080/08878730.2012.760024>
- Department of Basic Education (DBE). 2012a. Regulations pertaining to the national curriculum statement grades R-12. *Government Gazette*, 36041(1114).
- Department of Basic Education (DBE). 2012b. National protocol for assessment grades R-12. *Government Gazette*, 34600.
- Department of Higher Education and Training (DHET). 2011. Minimum requirements for teacher education. *Government Gazette*, 553(34467).

- Draeger, J., del Prado Hill, P., Hunter, L.R. & Mahler, R. 2013. The anatomy of academic rigor: The story of one institutional journey. *Innovative Higher Education*, 38(4), 267–279. <http://dx.doi.org/10.1007/s10755-012-9246-8>
- Draeger, J., del Prado Hill, P. & Mahler, R. 2014. Developing a student conception of academic rigor. *Innovative Higher Education*, 40(3), 215-228. <http://dx.doi.org/10.1007/s10755-014-9308-1>
- Gray, C. 2008. *Getting rigor right: Academic challenge without the backlash of failure*. Atlanta, GA: Southern Regional Education Board.
- Lincoln, M. 2010. *Academic rigour in science assessment tasks*. Available at www.eprints.qut.edu.au/33209/3/Mary_Lincoln_Thesis.pdf. [Accessed 2 May 2015].
- Mahler, R.E., Draeger, J., Del, P. & Hill, P. 2014. Comparing faculty and student conceptions of academic rigor. *The International Journal of Interdisciplinary Educational Studies*, 8(1), 31-41.
- McMillan, J. 2014. *Classroom assessment: Principles and practice for effective standards-based instruction*. London: Pearson.
- Reich, G.A., Turner, A.B. & Volkan, S. 2013. *Academic rigor for all: A review of literature*. Richmond, VA: Metropolitan Educational Research Consortium.
- Rudner, L. & Schafer, W.D. 2002. *What teachers need to know about assessment*. Washington, DC: National Education Association.
- Selkrig, M. & Keamy, R.K. 2015. Promoting a willingness to wonder: Moving from congenial to collegial conversations that encourage deep and critical reflection for teacher educators. *Teachers and Teaching: Theory and Practice*, 21(4), 421-436. <http://dx.doi.org/10.1080/13540602.2014.969104>
- Shalem, Y. & Rusznyak, L. 2013. Theory for teacher practice: A typology of application tasks in teacher education. *South African Journal of Higher Education*, 27(5), 1118–1134.
- Shalem, Y., Dison, L., Gennrich, T. & Nkambule, T. 2013. I don't understand everything here... I'm scared: Discontinuities as experienced by first-year education students in their encounters with assessment. *South African Journal of Higher Education*, 27(5), 1081–1098.
- Slonimsky, L. & Shalem, Y. 2006. Pedagogic responsiveness for academic depth. *Journal of Education*. 40:35 – 58.
- Stark, P.B. & Freishtat, R. 2014. *An evaluation of course evaluations*. Available at <http://www.stat.berkeley.edu/~stark/Preprints/evaluations14.pdf>. [Accessed 3 May 2015].
- Stiggins, R. 1991. Asssment literacy. *Phi Delta Kappa*, 72(7), 534–539.
- Stiggins, R. 1999. Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice*, 18(1), 23-27. <http://dx.doi.org/10.1111/j.1745-3992.1999.tb00004.x>