

2014

Title: How do Preservice Teacher Education students Move From Novice to Expert Assessors?

Peter R. Grainger

University of the Sunshine Coast, peter.grainger@usc.edu.au

Lenore Adie

Queensland University of Technology

Recommended Citation

Grainger, P. R., & Adie, L. (2014). Title: How do Preservice Teacher Education students Move From Novice to Expert Assessors?. *Australian Journal of Teacher Education*, 39(7).

Retrieved from <http://ro.ecu.edu.au/ajte/vol39/iss7/6>

This Journal Article is posted at Research Online.

<http://ro.ecu.edu.au/ajte/vol39/iss7/6>

How do preservice teacher education students move from novice to expert assessors?

Peter R. Grainger
University of the Sunshine Coast
Lenore Adie
Queensland University of Technology

Abstract: Despite the acknowledged importance of assessment in education, there has been minimal research into the preparation of preservice teachers for the important role of involving preservice teachers in marking, grading, moderating and providing feedback on student work. This article reports on a pilot project in which preservice teachers participated in an ongoing peer assessment and social moderation process in a dedicated course on assessment. The purpose of the project was to investigate specific ways in which key assessment processes can be effectively taught to preservice teachers. The research involved 96 preservice teachers who completed a Likert scale survey and free text responses to set questions. The results indicated that while preservice teachers valued the process, continual opportunities to learn the nature and purpose of essential assessment practices related to marking, grading, moderating and providing feedback are necessary to graduate competent and work-ready assessors.

Introduction

In Australia, preservice teacher education students gain teacher registration through completion of a recognised course of study which involves a mix of discipline courses and education-related courses including courses on curriculum, pedagogy and assessment. This range of study is considered necessary pre-requisite knowledge for teaching. A major part of teachers' work includes the assessment of students, evidenced by the National Professional Standards for Teachers in which Standard 5 requires teachers to "Assess, provide feedback and report on student learning" (Australian Institute for Teaching and School Leadership, 2011, p.16). Attached to this Standard are five sub-standards that require teachers to: assess student learning; provide feedback to students on their learning; make consistent and comparable judgements; interpret student data; and report on student achievement.

Key assessment skills involve marking, grading, moderating and providing feedback on student work. Despite the acknowledged significance of assessment, the exposure of preservice teachers to a range of specific assessment practices in their under-graduate courses has not been well investigated. The preservice teachers in this study were initially 'exposed' to these assessment practices via lectures and readings in the first four weeks of semester. This was followed by four weeks of authentic assessment activities, including grading of peers, the social moderation processes related to this grading, and peer feedback. The moderation discussion was focal to the preservice teachers' involvement in these assessment practices, and hence the learning that students may gain from involvement. Each of these assessment practices are discussed in the following section in terms of how they were addressed in the teacher education course.

The key research question was:

How do preservice teacher education students move from novice to expert assessors?

Related sub-questions were:

- *What do students know about assessment?*
- *How does moderation improve assessment knowledge?*

Learning to be an assessor in preservice teacher education courses

Despite the fact that assessing and accompanying participation in moderation processes is a significant part of a teacher's responsibilities, a review of the literature failed to reveal any reported research that investigated how preservice teachers learn about and are inducted into key assessment practices, including moderation, within their university coursework. While assessment is generally covered in preservice teacher education courses through dedicated courses and curriculum units, it is unclear if specific generic practices such as marking, moderation and providing feedback are covered. It is acknowledged that exposure to such key assessment practices often occurs in practicum experiences and as passive recipients in their own courses when assessment is done to students.

Learning to be an assessor through being assessed can be problematic, particularly when working between the different levels of education (for example, being assessed in higher education and becoming an assessor in a primary school setting). Personal experiences of being assessed can influence preservice teachers' beliefs and understandings of assessment which may propagate both good and poor practices. This idea links to the work of Lortie (1975) and then Borg (2004) who describe an apprenticeship of observation, specific to teacher education contexts during which students come to a tertiary course having already experienced thousands of hours of teacher practices (including assessment) and these experiences are responsible for many of their preconceptions. Thus, students are not in a position to analyse and evaluate these observed behaviours from a pedagogically sound perspective and hence they remain intuitive and imitative, described as 'folkways of teaching' (Lortie, 1975, p. 62).

To develop key assessment practices, we suggest in this article that preservice teachers need to be active participants in learning about assessment initially through traditional teaching and learning models such as lectures, and then through active involvement in purposeful assessment activities. Participation in moderation discussions involves the preservice teachers in making judgements using the provided criteria sheets as well as composing feedback to fellow students on their performance. These different key elements of assessment practice are outlined in the following sections.

Moderation

Moderation of assessment judgements involves teachers in matching evidence in a student work sample with a standard descriptor. The moderation practice is a significant activity for teachers as it is a form of quality assurance related to the consistency of their judgements of student work. Moderation is defined in this article as a practice of engagement in which teaching team members develop a shared understanding of assessment requirements, standards and the evidence that demonstrates differing qualities of performance. Its purpose is to ensure that there is consistency of judgements between assessors and that these judgements are aligned with established and visible criteria and standards in order to quality assure assessment processes (Adie, Lloyd & Beutel, 2013). Sadler (2009, p.2) uses the term 'consensus moderation' to describe the collaboration and discussion regarding the allocation

of marks and the establishment of 'a common view about the grading 'standards' to be used for the whole student group'. The Queensland Studies Authority (The State of Queensland, 2008) identifies three main moderation models: The Expert model, the Calibration model and the Conferencing model.

Moderation processes were 'taught' to the preservice teachers over four weeks through regular weekly two hour lectures. The three different moderation models (expert, calibration and conferencing) were presented to students during the semester using video segments of enacted moderation meetings (The State of Queensland, Queensland Studies Authority, 2008). The preservice teachers were able to critique the performances of moderation as presented in the videos, and identify specific behaviours that would enhance or inhibit a successful moderation meeting. The key inter-related messages regarding moderation that were presented included:

- Reaching consensus through rich conversations and professional dialogue
- Ensuring consistency of judgments through shared understandings
- Interpreting and applying standards in a common way
- Sharing and grading representative samples of student work across different standards
- Reaching an on-balance judgment of overall quality that takes into account achievement in different criteria across the assessment task.

Although institutionalised as accepted practice there appears to be limited understanding of moderation as an essential part of teaching and learning and significant confusion amongst academics in relation to shared understandings of criteria, standards and the qualities that provide evidence of a standard (Sadler, 2010). A number of solutions have been proposed including timing of the process (Bloxham 2009), training of assessors from day one of a course (Kuzich, Groves, O'Hare, & Pelliccione, 2010) and greater time allocated to students and staff to discuss standards (Bloxham, Boyd, & Orr, 2011). It is this last solution which inspired this current project. If preservice teachers receive their assessment training predominantly through their lived experiences in academia as undergraduate or postgraduate trainees and in teaching practice situations, it is vital that this assessment training provides quality models. Hence, in higher education lecturers and tutors need to become proficient in these assessment practices so that students are involved in an 'apprenticeship of observation' (Borg, 2004) that progresses their assessment practices.

The second phase involved the preservice teachers grading peer presentations during weekly tutorials from weeks 5-9. In this phase the preservice teachers were immersed in the moderation process, putting into practice the theoretical content learnt in phase one as active participants in negotiating grading decisions. The presentations were part of the formal assessment which required groups to present on an aspect of assessment. This assessment was graded by the tutor and also by peers. At the conclusion of the assessment, the presenting groups vacated the room and the peer audience engaged in an individual grading process using a criteria sheet. The class then broke into smaller groups to moderate their judgments made individually in order to reach a consensus on the overall on-balance judgement. Students were required to manage this process with no intervention from the tutor. The final group grade was given to the tutor who then debriefed the presenting group with the grade, and the specific feedback from each group. The peer feedback and awarded grade was used by the tutor to inform the final grade given to each presenting group. All feedback and criteria sheets were de-identified to preserve anonymity.

Criteria sheets

Criteria sheets typically contain three elements: criteria, standards and standards descriptors. The description of quality identified in each of the standards is matched against evidence in student work. Sadler (1987) noted that descriptions of standards are often vague, unclear, indicative only, open to interpretation and make assumptions that the user is familiar with the language used in the descriptions. If the description of quality is not explicit then the matching exercise becomes problematic due to different interpretations by different assessors of the quality described. This results in inconsistency in grading, and in the quality of feedback provided to students.

One of the aims of this project was to expose students to the use of criteria sheets and more specifically, conversations around standards descriptors, as the fundamental tool used in assessing student work, and as an integral aspect of the moderation process aimed at achieving consistency of teacher judgements. In order to be competent assessors upon graduation, students clearly require training in using criteria sheets.

During lectures and tutorials, the preservice teachers were exposed to two styles of criteria sheets: the traditional matrix style criteria sheet and the Continua model of a Guide to Making Judgements. The latter model was a model used by the Queensland Studies Authority to assess Queensland Comparable Assessment Tasks (QCATs) introduced in 2009 under the Queensland Curriculum, Assessment and Reporting Framework (QCARF). Students viewed, discussed and critiqued authentic exemplars of these models. It was decided that the preservice teachers would use the continua style of criteria sheet to make their judgements during the presentations (see Appendix 4).

Feedback

The final key assessment practice focussed on in the course was feedback. Feedback on student work is essential to student learning and effective comments on student work are considered an aspect of quality teaching practice (Hattie, Biggs & Purdie, 1996; Black & Wiliam, 1998; Ramsden, 2003). Despite its acknowledged importance in the learning loop, feedback is not a concept that has been adequately explored (Carless, 2006). Hence, this assessment process was considered a major focus for this pilot study. The core messages regarding feedback shared with students included the purpose of feedback, how to provide quality feedback, and different modes of feedback including electronic, oral and written. Feedback was considered an important skill to explore thoroughly with students since research into student experience of feedback throughout their education, including their university study, has shown that while feedback communications may be appreciated by students, they often lead to little if any improvement in subsequent submissions (Sadler, 2009). One of the major obstacles to ensuring that feedback is effective is the difference between student and marker perceptions of feedback. Carless (2006) found that there were marked differences in the way that students and markers perceived feedback.

Sadler (2009) claims that for feedback to be effective, the feedback statements made by the assessor must be able to be understood by students. Misunderstanding can occur when an assessor makes assumptions that a student understands the concepts and terms that are described routinely by the assessor.

In phase one of the study, the preservice teachers discussed and analysed topics such as the use of criteria sheets to provide feedback, frameworks for providing feedback, developing peer feedback, critical evidence, and sharing learning intentions and success criteria to promote their students' understanding of feedback. In phase two, the preservice teachers practiced these skills by providing feedback to their peers through written comments

on their presentations and completion of the criteria sheet. The students followed a framework for providing effective feedback that involved them identifying aspects of the presentation that were done well, aspects to further develop and advice on how to develop these skills. This framework was based on advice provided by Black and Wiliam (1998) on effective feedback, and engaged the students in linking feedback with strategies for improvement.

Method

A social moderation process was implemented in a mixed cohort preservice teacher education course (under-graduate and graduate diploma) dedicated to assessment and involving a total of 130 enrolled students. In addition to other assessment principles, students had been exposed to social moderation concepts in lectures leading up to the social moderation process in tutorials, which commenced half way through the course. Over a period of five consecutive weeks, students participated voluntarily in a social moderation process involving their own assessment of a peer group’s summative assessment performance during tutorials. At the end of each performance, students peer assessed the performance in small groups in order to reach consensus and grade the performance. This peer assessment was later provided to the performing group. The peer assessment feedback was anonymous. Performing groups vacated the tutorial room during the moderation process.

At the end of the semester, students were asked to complete a hard copy survey related to the process (refer Appendix one). The survey was a mix of Likert scale responses and free text comments. These were collected and analysed. 96 surveys were returned. Students completed only one survey response and hence this represents a 74% response rate. The survey included three sections, a demographics section, a Likert scale response section and a free text section. The ages of students ranged from 18-63. There were 51 males and 55 females. All were from a secondary specialisation. These demographic details are represented in Table 1.

There are two limitations to this study. The major limitation is the small sample size, only 96 respondents. Further studies of this kind with larger samples will ensure validity of the findings. In addition, one-on-one and/or focus group interviews with selected students would enable a greater depth of understanding to be reached in terms of the research questions asked.

Age	Unknown	Less than 20 yrs	20-30 yrs	30-40 yrs	40-50 yrs	50+ yrs
	5	7	61	15	7	1
Gender	Male	Female				
	41	55				
Program	BEd	Grad Dip	Unknown			
	26	58	12			

Table 1: Demographics of preservice teacher participants

Results

Section Two results (Likert scale responses): Analysis of the responses and the relative percentages for each of the 21 questions indicated that students perceived that their learning about moderation and assessment had been supported as a result of their participation in the moderation discussion (question 2: 82%; question 8: 61%). Students indicated that they regarded the social moderation process as important for making consistent judgements of student work (question 9: 78%). Despite acknowledgment of the importance of reaching a consensus about 30% of students noted that they did not reach a consensus (question 16 and 17); 76% felt reaching consensus was difficult (question 18), yet 83% noted they were satisfied with the final judgement (question 20). Preparation for the activity in lectures leading up to the activity was also considered successful as a majority of students felt prepared for the moderation activity (question 7: 69%; question 11: 69%; question 12: 68%). These positive results are despite the fact that only 27% were familiar with the Continua model of a criteria sheet (question 1) and only 37% indicated that they had confidence in their assessment ability prior to the moderation activity (question 4). Domination by peers was a significant issue with 56 and 57% of students (question 13 and 15). The response percentages to each of the 21 questions are identified below.

1. 57% were not familiar with the criteria sheet model; 27 % were
2. 82% thought the moderation process of grading peers helped to better understand assessment
3. 45% were initially nervous about grading peers
4. 37% were confident with knowledge of assessment prior to the experiences; 30% were not; 29% were unsure
5. 76% disagreed with the statement that grading peers was unfair
6. 62% had graded an assessment task prior to this experience; 31% had not
7. 20% felt unprepared for the grading task; 69% felt prepared
8. 61% felt better prepared for a role as assessor upon graduation; 12% did not; 20% unsure
9. 78% thought discussing grading decisions with peers was crucial to reaching consensus
10. 64% thought the grading tool used was explicit enough to make accurate grading decisions
11. 69% understood what evidence to look for in student work as a result of the criteria sheet
12. 68% knew where to begin the grading process
13. 56% felt some peers dominated the moderation discussions
14. 70% felt confident in expressing grading decisions
15. 57% felt intimidated by some peers during moderation discussions
16. 69% reached a consensus about the performance under review
17. 71% reached a consensus about the performance under review
18. 76% thought reaching consensus was a difficult process
19. 56% thought reaching a consensus was an easy process
20. 83% were happy with the final on balance judgement
21. 76% knew enough about the topic to make an accurate decision alone

The 'raw' response rates to the above questions are identified in Appendix 2.

Section three (free text) results: The free text responses to three questions were analysed for consistent patterns resulting in a number of key themes that emerged from each of the three questions in this section. The three questions focussed on identifying the most valuable part of the process, the least valuable part of the process and recommendations that would improve the student experience. These themes were confirmed independently by a colleague. These themes are represented in Table 2 below.

Question	Key Theme
Valued	<ul style="list-style-type: none"> • Gaining experience on giving feedback • Doing assessment-learning how to mark/grade/assess • Learning how to assess from peers through discussion, reflection and justifying decisions • Increasing confidence in grading /assessing using criteria-others also valued what I valued • Having second opinions to confirm my judgements/decisions • Arriving at consensus through discussion • Gaining peer perspectives about what was valued in case I missed something • Using criteria to explain/justify a result.
Not valued	<ul style="list-style-type: none"> • Having to reach consensus • Not understanding the criteria • Some people dominated discussions • Not seeing all the presentations first before grading • Not knowing enough about assessment, not enough experience • Not knowing what the expert teacher thought and being able to compare • Not having an official voice in the formal result • Having an official voice in the formal result.
Recommended	<ul style="list-style-type: none"> • Preparing us better on how to give constructive feedback, how to assess, how to moderate • Not having to reach a consensus • Understanding the language of the criteria sheet, more explanation needed • Unbiased discussion • More practice of this process • More detailed criteria sheet • More conversation on what each level looks like, what the standards descriptors mean • Using marks (not just standards); having explicit weightings • Changing the grading for the first group as later groups are advantaged • Students discussing feedback directly with other students after the presentation.

Table 2: Key themes

Discussion

The following discussion focuses on key findings from the survey and free text results: 1) while preservice teachers noted their increased knowledge about moderation practices and judgement making, they were frequently unable to reach consensus on the final grade, 2) while the majority of preservice teachers felt that they were better prepared for their role as assessors as a result of this experience, a significant minority still indicated a lack of knowledge, 3) the value of the criteria sheet to make a judgement and provide detailed useful feedback remains unclear, and 4) being an assessor in a standards-referenced assessment system involves a complex interplay that require practice to learn well.

Working in a system based on standards-referenced assessment requires an epistemological shift in thinking about curriculum, assessment and the positioning of the student in the learning process. Involving preservice teachers in the negotiated practice of making judgements introduced them to this way of thinking and working, but the results from this study show that this immersion through a single course, on its own, is insufficient. This is evidenced in the percentage of students who acknowledged the degree of difficulty in

reaching consensus and the percentage who remained not confident with their grading decisions. The students' recommendations suggest that they had not yet developed a full understanding of standards-referenced assessment as involving shared understanding of a standard in order to reach consensus. Specifically, their request for more detailed criteria indicates that the philosophical basis of a standards-referenced curriculum needs much more time to develop. Lave and Wenger (1991) highlight the complex process of moving from novice to expert in a community of practice that includes the influence of historic practices as well as the multiple opportunities for learning offered through the community.

Analysis of the results reveals that preservice teachers perceived that their participation in the process of social moderation with peers enhanced their knowledge of assessment practices. These discussions with peers served to allay a number of fears in relation to the inexperience of students in terms of marking/grading performance. The notion of novice assessor is evidenced in the reactions of many students who indicated a deep appreciation of their own limitations in grading work and the significant advantages of dissecting criteria, discussing what is valued in an assessment performance and finally arriving at a consensus as a result of these discussions. An advantage of such a process was the fact that multiple assessments by different group members ensured that all aspects of performance of the group being assessed were scrutinized. For example, some students focussed on content knowledge, others focussed on engagement in activities, when deciding an overall on-balance judgement of performance; despite the fact that criteria were equally weighted. Others interpreted the standards descriptors differently and valued different aspects of the criterion over others. Discussing these differences and interpretations was valuable, in fact enlightening as these aspects of performance had not initially been considered by the students when grading the work individually prior to the moderation discussions.

In terms of hindrances to learning about moderation practices analysis of the results revealed misunderstandings of the purpose of assessment moderation through consensus, evidenced by opposing views by different groups of students in relation to having to reach a consensus. Some students highlighted the need to reach a consensus as a major issue. Thus, the findings provide evidence that the very concept of consensus moderation was not understood by all students, despite this focus in the lead-in lectures and tutorials. Possibly the message delivered was not interpreted correctly, possibly students did not attend the lectures and possibly students had not engaged with the readings. Additionally, many responses indicated that moderation discussions were sometimes dominated by personalities that ensured that some opinions were not heard or not able to be voiced. This occurred despite the fact that students had been exposed to the required protocols in lead-in lectures through a QSA produced DVD that included role plays modelling good and bad social moderation practices. The domination of peers in the moderation process suggests that further and stronger reinforcement of protocols around these processes is required possibly through the preservice teachers' participation in role play activities during tutorials.

There were conflicting opinions on the level of preparation of students which was identified as a significant concern and many students (20%) indicated they felt under-prepared for the activity despite the fact that this activity had been preceded by four weeks of lectures dedicated to assessment principles including moderation processes. The difference between peripheral participation of a practice through hearing about it and observing the practice, and active participation in the practice that involves working through the complexities of negotiated consensus-making clearly shows the need for continual learning opportunities. Lave and Wenger (1991) demonstrated how practice develops over time as different skills are learnt within a holistic notion of the wider practice as experienced by experts.

Another significant hindrance identified by students related to the criteria sheet and the vagueness of the standards descriptors used, which required much discussion and dissecting in order for all participants to have a common understanding. Many students requested that more detailed criteria sheets be provided. Sadler (2007, p. 390) has described the over specification of criteria as reducing learning to "pea-sized bits...swallowed one at a time" which shifts the focus from a holistic notion of learning to one of discrete elements that students may or may not be able to piece together. Designing criteria that clearly identify the quality of work yet allow for multiple ways to evidence this quality is considered one of the strengths of a standards-referenced assessment system. This illustrates the necessity of having moderation processes to develop shared understandings of the evidence that denotes the quality of a response.

A revealing theme was the requirement or desire of some students to know how the lecturer as the acknowledged expert had graded the performances in order to validate or even calibrate their own judgements. When learning new practices, novices look to experts for guidance. Observations of experts at work and conversations with experts enable novices to assimilate new practices into their repertoire. In this project, students were able to view expert moderators at work by viewing video segments of moderation discussions. In the absence of an expert facilitator who is experienced and can guide the moderation discussions, the preservice teachers may benefit from the provision of questioning frameworks that provide guidance of the type of probing questions that support deeper analysis of judgement decisions.

A final hindrance for some students working in a system of standards-referenced assessment was their belief that they needed to see all performances prior to assigning grades. This belief indicates their misunderstanding of the criteria and standards-referenced approach and is reflective of a norm-referenced system which determines results based on comparisons with other group performances. While the preservice teachers acknowledged their learning about assessment through immersion in these experiences, such comments clearly evidence their status as novices in standards-referenced assessment practice. It is clear from this result that further teaching of the moderation process and how to judge against standards descriptors that includes opportunities for dialogue between expert and novice assessors and practice assessment sessions are required before we immerse preservice teachers directly in the process of judging and moderating assessment performances by peers. In short a scaffolded model is required that is developed over a course of study rather than just one subject.

Conclusions

Despite the acknowledged importance of assessment in tertiary education courses, there are no published research findings related to the preparation of preservice teacher educators for the important process of marking, grading, moderating and providing feedback on student work, upon graduation. In addition, few universities offer dedicated assessment courses, resulting in an indication that teacher preparation courses graduate novice assessors, generally untrained in the important routine teacher tasks related to assessment. This article has reported on a pilot project aimed at investigating preservice teacher responses to an ongoing peer assessment and social moderation process in a dedicated course on assessment. The results of this pilot project are promising and serve as validation of the project and encouragement to further investigate ways in which teacher preparation courses can successfully train novice assessors. It can be concluded that active and ongoing participation in peer-assessment activities and associated moderation processes is a necessity

to graduate competent assessors. The key findings suggest that students perceived the process as valuable, assessment knowledge was significantly increased and confidence in grading processes was enhanced. However, results of this pilot project indicate that despite being exposed directly to the principles of assessment during regular coursework leading in to the peer assessment and moderation activities, some students were still confused as to the fundamental tenets of work in a standards-referenced assessment system which involves developing shared understandings of the qualities that provide evidence of a standard. The implications of these results suggest a need for more a scaffolded and developmental preparation prior to preservice teacher engagement in peer assessment grading and moderation processes.

References

- Adie, L., Lloyd, M. & Beutel, D. (2013). Identifying discourses of moderation in higher education. *Assessment & Evaluation in Higher Education*, 38(8), 968-977.
<http://dx.doi.org/10.1080/02602938.2013.769200>
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.
- Bloxham, S. (2009). Marking and moderation in the UK: False assumptions and wasted resources. *Assessment & Evaluation in Higher Education*, 34 (2), 209-220.
- Bloxham, S., Boyd, P. & Orr, S. (2011). Mark my words: The role of assessment criteria in UK Higher Education grading practices. *Studies in Higher Education*, 36 (6), 655-670.
- Borg, M. (2004). The apprenticeship of observation. *ELT Journal*, 58(3), 274-276.
- Carless, D. (2006): Differing perceptions in the feedback process. *Studies in Higher Education*, 31(2), 219-233. <http://dx.doi.org/10.1080/03075070600572132>
- Hattie, J., Biggs, J. & Purdie, N. (1996) Effects of learning skills intervention on student learning: a meta-analysis. *International Journal of Educational Research*, 11, 187-212.
- Kuzich, S., R. Groves, S. O'Hare, & L. Pelliccione. (2010). *Building team capacity: Sustaining quality in assessment and moderation practices in a fully online unit*. Paper presented at the ATN Assessment Conference, University of Technology Sydney, Australia. Retrieved from <http://www.iml.uts.edu.au/pdfs/Kuzich.pdf>.
- Lave, J., & E. Wenger. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.
<http://dx.doi.org/10.1017/CBO9780511815355>
- Lortie, D. (1975). *Schoolteacher: A sociological study*. London: University of Chicago Press.
- National Professional Standards for Teachers (2011). Education Services Australia as the legal entity for the Ministerial Council for Education, Early Childhood Development and Youth Affairs (MCEECDYA). pp 1-28.
- Ramsden, P. (2003). *Learning to teach in higher education* (2nd ed.). London, Routledge Falmer.
- Sadler, D. R. (1987). Specifying and promulgating achievement standards. *Review of Education*, 13, 191-209. <http://dx.doi.org/10.1080/0305498870130207>
- Sadler, D. R. (2007). Perils in the meticulous specification of goals and assessment criteria. *Assessment in Education: Principles, Policy and Practice*, 14(3), 387-392.
<http://dx.doi.org/10.1080/09695940701592097>
- Sadler, D.R. (2009). Indeterminacy in the use of preset criteria for assessment and grading in higher education. *Assessment and Evaluation in Higher Education*, 34, 159-179.
<http://dx.doi.org/10.1080/02602930801956059>
- Sadler, D.R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment and Evaluation in Higher Education*, 35, 535-550.
<http://dx.doi.org/10.1080/02602930903541015>
- The State of Queensland (Queensland Studies Authority) (2008). *Building Student Success*. Retrieved from
http://www.google.com.au/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=2&ved=0CDoQFjAB&url=http%3A%2F%2Fwww.qsa.qld.edu.au%2Fdownloads%2Fp_10%2Fqcar_building_student_success.pdf&ei=CtJcUp_tO6SliQfFz4HQBA&usq=AFQjCNEiatVLXvlpLoP7EqwwcRtpQFeEYQ&sig2=urxtSHETi09FM9EDFG7H9w&bv m=bv.53899372,d.dGI

Appendix 1: Survey to students re moderation processes

Part A: Demographics

Age:

Gender: Male or Female

Program: Bachelor of Education or Graduate Diploma of Education

Part B: Survey: Circle the most accurate response for each statement.

I was familiar with the Continua model of a Guide to Making Judgements prior to the experiences

Strongly agree **Agree** **Unsure** **Disagree** **Strongly disagree**

The moderation process of grading peers helped me to better understand assessment

Strongly agree **Agree** **Unsure** **Disagree** **Strongly disagree**

I was initially nervous about grading my peers

Strongly agree **Agree** **Unsure** **Disagree** **Strongly disagree**

I was confident with my knowledge of assessment prior to the experiences

Strongly agree **Agree** **Unsure** **Disagree** **Strongly disagree**

Grading peers was unfair

Strongly agree **Agree** **Unsure** **Disagree** **Strongly disagree**

I had never graded an assessment task prior to this experience

Strongly agree **Agree** **Unsure** **Disagree** **Strongly disagree**

I felt unprepared for the grading task

Strongly agree **Agree** **Unsure** **Disagree** **Strongly disagree**

As a result of my experiences I feel better prepared for my role as assessor upon graduation

Strongly agree disagree	Agree	Unsure	Disagree	Strongly
------------------------------------	--------------	---------------	-----------------	-----------------

Discussing grading decisions with peers was crucial to reaching consensus

Strongly agree disagree	Agree	Unsure	Disagree	Strongly
------------------------------------	--------------	---------------	-----------------	-----------------

The grading tool used was explicit enough to enable me to make accurate grading decisions

Strongly agree disagree	Agree	Unsure	Disagree	Strongly
------------------------------------	--------------	---------------	-----------------	-----------------

As a result of the standard descriptors I understood what evidence to look for in student work

Strongly agree disagree	Agree	Unsure	Disagree	Strongly
------------------------------------	--------------	---------------	-----------------	-----------------

I did not know where to begin the grading process

Strongly agree disagree	Agree	Unsure	Disagree	Strongly
------------------------------------	--------------	---------------	-----------------	-----------------

Some peers dominated the moderation discussions

Strongly agree disagree	Agree	Unsure	Disagree	Strongly
------------------------------------	--------------	---------------	-----------------	-----------------

I felt confident in expressing my grading decisions

Strongly agree disagree	Agree	Unsure	Disagree	Strongly
------------------------------------	--------------	---------------	-----------------	-----------------

I felt intimidated by some of my peers during moderation discussions

Strongly agree disagree	Agree	Unsure	Disagree	Strongly
------------------------------------	--------------	---------------	-----------------	-----------------

My group reached a consensus about the performance under review

Strongly agree disagree	Agree	Unsure	Disagree	Strongly
------------------------------------	--------------	---------------	-----------------	-----------------

My group did not reach a consensus about the performance under review

Strongly agree disagree	Agree	Unsure	Disagree	Strongly
------------------------------------	--------------	---------------	-----------------	-----------------

Reaching consensus was a difficult process

Strongly agree **Agree** **Unsure** **Disagree** **Strongly disagree**

Reaching a consensus was an easy process

Strongly agree **Agree** **Unsure** **Disagree** **Strongly disagree**

Although my group reached a consensus I was not happy with the final on balance judgement

Strongly agree **Agree** **Unsure** **Disagree** **Strongly disagree**

I felt I did not know enough about the topic to make an accurate decision alone

Strongly agree **Agree** **Unsure** **Disagree** **Strongly disagree**

Part C: Free text

The most valuable part of the process was:

The least valuable part of the process was:

To enable a better experience for future students I would recommend:

Appendix 2: Likert Scale responses (n=96)

Q	Strongly agree	Agree	Unsure	Disagree	Strongly disagree
1	4	23	12	32	25
2	10	72	8	5	1
3	6	39	12	31	8
4	4	33	29	25	5
5	0	7	13	61	15
6	6	25	3	45	17
7	1	19	7	58	11
8	6	55	20	11	1
9	17	61	9	7	2
10	11	53	16	13	3
11	9	60	12	12	3
12	2	22	4	57	11
13	4	29	7	54	2
14	11	59	9	15	2
15	2	30	7	46	11
16	16	53	2	22	3
17	7	14	4	50	21
18	1	13	6	61	15
19	7	49	8	30	2
20	1	8	4	69	14
21	2	10	8	68	8

Appendix 3: The response percentages to each of the 21 questions

1. 57% were not familiar with the criteria sheet model; 27 % were
2. 82% thought the moderation process of grading peers helped to better understand assessment
3. 45% were initially nervous about grading peers;
4. 37% were confident with knowledge of assessment prior to the experiences; 30% were not; 29% were unsure
5. 76% disagreed with the statement that grading peers was unfair
6. 62% had graded an assessment task prior to this experience; 31% had not
7. 20% felt unprepared for the grading task; 69% felt prepared
8. 61% felt better prepared for a role as assessor upon graduation; 12% did not; 20% unsure
9. 78% thought discussing grading decisions with peers was crucial to reaching consensus
10. 64% thought the grading tool used was explicit enough to make accurate grading decisions
11. 69% understood what evidence to look for in student work as a result of the criteria sheet
12. 68% knew where to begin the grading process
13. 56% felt some peers dominated the moderation discussions
14. 70% felt confident in expressing grading decisions
15. 57% felt intimidated by some peers during moderation discussions
16. 69% reached a consensus about the performance under review
17. 71% reached a consensus about the performance under review
18. 76% thought reaching consensus was a difficult process
19. 56% thought reaching a consensus was an easy process
20. 83% were happy with the final on balance judgement
21. 76% knew enough about the topic to make an accurate decision alone

Appendix 4: Sample Continua model of a Guide to Making Judgements (source: authors)

Task specific criterion	Task specific criterion	Task specific criterion	
Standards descriptor for A	Standards descriptor for A	Standards descriptor for A	A
Standards descriptor for B	Standards descriptor for B	Standards descriptor for B	B
Standards descriptor for C	Standards descriptor for C	Standards descriptor for C	C
Standards descriptor for D	Standards descriptor for D	Standards descriptor for D	D
Standards descriptor for E	Standards descriptor for E	Standards descriptor for E	E