Large class teaching: how does one go about the task of moderating large volumes of assessment?

Dr Jaclyn Broadbent¹,²,³

¹ School of Psychology, Deakin University

² CRADLE Fellow at Deakin's Centre for Research in Assessment and Digital Learning

³ Corresponding author:
Dr Jaclyn Broadbent
School of Psychology, Deakin University,
221 Burwood Hwy, Burwood 3125, Victoria, Australia.
Phone: +61 3 9244 3043
Fax: +61 3 9244 6858
Email: jaclyn.broadbent@deakin.edu.au

Recommended Citation


This is a post-peer-review, pre-copyedit version of an article published in Active Learning in Higher Education. The final authenticated version is available online at: https://doi.org/10.1177/1469787417721360

Conflict of Interest: The authors declare that they have no conflict of interest.
Abstract
There is limited research on the quality of assessment moderation in large classes. Effective moderation practices can improve reliability, as well as reduce marker bias, attenuate prevalence of ‘hard’ or ‘soft’ markers, increase student and staff confidence in marking, and enhance the development of staff. This article shares a marking moderation practice used in large class teaching (>1500 students). The article highlights the importance of: (1) including resources / communication that are provided to markers in order to facilitate a shared understanding and interpretation of the marking criteria; (2) incorporating multiple points of double marking to detect differences in marker performance over time that may have been brought on by fatigue, tight timelines, or inexperience; and (3) developing markers’ skills early through formative feedback to acquire self-sufficiency, accuracy, and expertise in the grading process. The practice of moderation utilised in this article overcomes some of the challenges of moderating large volumes of assessments. Particularly, the use of audio feedback, and video resources were the most novel and useful.

Key words: moderation; quality assurance; assessment; higher education; large classes; audio feedback; grades

Quality assurance is a key driver in rethinking curriculum design for tertiary courses in the higher education context. This issue has been brought to the forefront by, for example, the higher education accreditation body Tertiary Education Quality and Standards Agency (TEQSA), who have increased emphasis on the reliability and validity of marks given for assessment (Kuzich et al., 2010). Accordingly, universities in most countries, including Australia, are now required to declare details of their moderation processes, noting differences in these processes across delivery methods, sites, and student cohorts for quality assurance purposes (TEQSA, 2013).

Moderation itself is “a process for assuring that an assessment outcome is valid, fair and reliable, and that marking criteria have been applied consistently” (Bloxham 2009, p.4). Ensuring an effective moderation process can improve reliability, as well as reduce marker bias, attenuate ‘hard’ or ‘soft’ markers, increase student and staff confidence in marking, and enhance the development of staff (Bloxham, 2009). Common moderation practices include applying assessment criteria, assigning marks / grades, social moderation (discussing and negotiating assigned grades), and double marking student work (Klenowski et al., 2011). However, the increasing class sizes and staff workload in the higher education sector (Larkins, 2011) threatens academics’ capacity to provide and manage effective moderation processes and practices, giving rise to concern about the quality of the feedback given to students (Mutch, 2003). Further, how each marker goes about the task of marking is often not well known (Yorke et al., 2000). Thus, it is important to recognise the amount of time required for a marking team to reach a shared understanding and interpretation of assessment criteria and apply consistent marking. Producing high-quality feedback is a time-consuming and labour-intensive task for markers (Jones and Gorra, 2013), while short turn-around times for assessments remain unchanged by the University. Nevertheless, quality assurance must be maintained.

When dealing with large classes (for example, 1500+ students), the issue of quality assurance becomes more prominent and challenging. This is particularly the case when subject coordinators are solely responsible for implementing and overseeing assessment moderation. It is even more challenging when working with different markers across different campuses, where students are enrolled in a variety of study modes (on-campus, off-campus and blended study modes). Large student numbers are often accompanied by sizeable marking loads (Jessop et al., 2012), which necessitates reliance on sessional marking staff, to help assess all assignments. In Australia, for example, 80% of all undergraduate first-year marking and teaching is completed by sessional staff (Percy et al., 2008). Skill set, level of education, expertise, and experience often varies between markers, with some staff, including sessional staff, having limited prior marking experience. As such, ensuring consistency of marks and feedback, ensuring efficiency, and avoiding ‘disasters’ is a major challenge for the subject coordinator. Given these challenges, it is important to adapt moderation practices to scale. Based on the literature, several important aspects of moderation have been identified to enhance the moderation process in this context. These include (but are not limited to): shared understanding of the marking requirements, use of
blind double marking, and continuous marker development through feedback (Bloxham et al., 2015a; Grainger et al., 2008; Hand and Clewes, 2000).

Important to the moderation process is the development of resources for, and open communication with, the staff who are marking the students’ work. Grainger et al. (2008) argue that for the quality of student work to be assessed accurately, markers must have a strong understanding of the marking criteria and the quality associated with each grade level. While an accompanying rubric allows marking judgments to be both more systematic and transparent to the student, there is still leeway for varying interpretations of the criteria amongst marking staff (Andie et al., 2013). Kuzich et al. (2010) argue that large cohorts and a large number of markers exacerbate these problems, and therefore raise potential issues of maintaining quality assurance.

As such, sharing and discussion of the assessment rationale with markers as well as discussion about the expectation of marking staff in relation to the marking criteria is important in a large class situation. Taras and Davies (2012) found that lecturers did not necessarily have a ‘clear cogent, coherent and shared’ (p. 58) understanding of assessment, and that possibly assessment, and its consequences (for example, grading) were not thought about in sufficient depth. It is necessary therefore for the subject coordinator to provide resources that facilitate the marker’s understanding of both these requirements. This process also provides an opportunity for markers to seek clarification from the subject coordinator on any issues to do with processes or procedures in the assessment phase. According to Grainger et al. (2008), markers can only evaluate the quality of the assessment if they have a sound understanding of what represents ‘quality’ for any given assessment task. Ensuring that markers understand the assessment rationale and discuss what represents indicators of quality aids in this process.

To ensure that markers are marking to the same standard and have the same shared understanding, blind double marking can be employed. While double marking is said to be ineffective when used as the sole moderation practice (Yorke, 2003), it can be a useful device to help markers and subject coordinators compare marking across groups (Hand and Clewes, 2000). However, there are several potential pitfalls of double marking that need to be guarded against. Double marking can be problematic when the subject coordinator knows the original mark, if only fail assessments are double marked, if double marking occurs at the end of the marking process (Partington, 1994) and prevents students from receiving timely feedback (Jessop et al., 2012), or when it is assumed that a small sample of papers is representative of the whole cohort (Bloxham, 2009).

However, double blind-marking helps to avoid the potential problem of convergence towards the mark originally awarded, particularly when the first marker is experienced (Hand and Clewes, 2000; Partington, 1994). The double blind marking process also increases fairness in marking (Bloxham, 2009). Lastly, blind double marking early in the marking timeline helps to provide just-in-time feedback to markers to allow them to recalibrate and reduce any mistaken beliefs in applying assessment criteria to the task being assessed.

Recalibration actually helps develop the skills of the marker by giving them formative feedback on their assessment work. High quality and timely feedback on assessment is crucial for improving student knowledge and skill acquisition (Boud and Molloy, 2013; Shute, 2007). The same principles can be applied to staff. Formative feedback can:

- Clarify what appropriate performance looks like and how well one is progressing (Nicol and Macfarlane-Dick, 2006),
- Correct and prevent inappropriate task strategies or procedural errors (Shute, 2007), and
- Improve one’s learning and critical thinking skills (Fluckiger et al., 2010).

Assessment feedback during the semester is crucial for addressing discrepancies between learner and teacher perspectives, and for clear communication of expectations and standards of performance (Boud and Molloy, 2013). Marking should be transparent, and markers / coordinators should refrain from post hoc changes in the criteria, which in turn changes the marks being allocated (Yorke et al., 2000). Rust et al. (2005) argue that assessment needs to be approached in the same way as learning, where markers and coordinators create a feedback loop.
Creating a feedback loop to enable the markers to judge how their marking compares to others can also aid moderation. It is important to understand, however, that while the distribution of marks can provide insight into the standard being applied by all markers, it should not be the sole factor in determining whether marks are adjusted (Andie et al., 2013). Caution must be given to prevent defensive marking, where the marker becomes risk adverse and errs towards the cohort average (Hornby, 2003). If these factors are considered, it does give the marker a guide, particularly if their marking is more than one standard deviation above or below the marker cohort average.

Based on the above information, there is a need to develop moderation practices in a large class setting, which does the following three things. One, increase communication and shared understanding of the quality of students’ work. Two, commence the moderation prior to marking, and continue throughout the marking process. Three, develop marker expertise. The practices need to guide the moderation process so that they provide moderation in an effective and timely manner, and are manageable and produce quality moderation outcomes. With these imperatives in mind, this article describes a moderation practice and reports on its use in in a setting with large student numbers. The moderation practice described focuses on (1) good communication and resources with and for markers, (2) blind double marking early in the marking process (and randomly in the later stages), and (3) on-going development of marker expertise.

Method

Context

Deakin University, Australia has over 2100 students enrolled annually in a first year core Health Psychology subject. Its largest intake, Semester 2, has over 1500 students enrolled at the same time. Students are primarily from courses in the Health Faculty, although students from other Faculties also take the subject. As part of the assessment, students complete three reflective journal entries over the course of the 11-week semester, and receive a hybrid of formative and summative audio feedback from their marker at three different time points. The aim is to provide students with feedback for the current piece of work, while also providing feed-forward feedback. Feed-forward refers to ‘timely and constructive feedback’ that is useful for the next assignment (Hounsell et al., 2008, p65). Aside from audio feedback from markers, grading also includes a marking rubric and a generic comment linking the assessment to the learning outcomes of the subject. In Semester 2, the largest semester in 2014, the 1500+ students collectively submit around 4,500 reflective journal entries to be marked.

The marking period for each assessment is typically 18 days, and feedback and grades are available to students within 20 days. Students were given assignment return dates at the start of semester. Approximately 20 sessional marking staff were employed to complete marking in the large semester. All of the sessional staff were currently studying, in either a Masters, Doctorate or PhD in Psychology. There are two coordinators (full-time academic staff) who moderate the 20 sessional staff, many of whom they have never met in person. The subject runs across three campuses (hundreds of kilometers apart) and runs in a blended mode (combination of online and face-to-face) as well as online-only mode; factors that contribute to increasing the challenge of providing consistent, equitable, and accurate assessment feedback. While the processes described below has evolved over time since 2011, for the purpose of this article only the semester 2, 2014 period of assessment will be focused on. Each of the steps in this process is discussed in turn.

Good communication and resources with and for markers

The first step in the proposed moderation practices for large classes is based on Granger et al.’s (2008) argument that, for the quality of student work to be assessed accurately, markers must have a strong understanding of the marking criteria and the quality associated with each grade level. Below are examples of the key types of resources / communication that are provided to markers in order to facilitate a shared understanding and interpretation of the marking criteria. They are shared prior to the commencement of marking, as it provides an opportunity for markers to seek clarification early. Given the constraints of large marker numbers and their disparate geographical location, these resources are provided in a shared online space.

1. Marking guide. Includes assignment outline information, copies of marking rubrics, marking start and due dates and clear expectations about what is required from them as markers. Further, many of the below examples are found within the marking guide.
2. Videos discussing criteria within the marking rubric. Three videos were created discussing the marking criteria of each of the assessments. The subject coordinators discuss the knowledge/learning that each criterion captures, while giving concrete examples of student work that represent different levels of the criteria. For example, ‘To get a credit for this criterion, a student may have said something like X, Y, or Z to answer the question. They most likely will have left out W in their discussion.’ Links for the videos were provided in the marking guide.

3. Video explaining feed-forward feedback. One video was created of the coordinator (author) explaining what feed-forward feedback is, and giving examples of how it can be used in feedback to students. The link was provided in the marking guide.

4. Feedback examples. As many of the markers have never given audio feedback to students before, written examples illustrating the comments for different criteria levels were provided. We provided examples of different responses, appropriate language, level of encouragement, and degree of detail that is expected. Importantly, they are details about how to give feed-forward feedback to students, with concrete examples of how responses between journal assessments link together. These can be read out as a script, if required. Examples were provided within the marking guide.

5. Marking meetings. Before each marking period a 1-hr face-to-face meeting with videoconferencing was run. Meeting covered information in the marking guide and allowed time for markers to ask questions. This was also recorded.

6. Recording of the tutorial class that discusses the assessment piece. To ensure that non-teaching markers were familiar with the content that related to the assessment, they were provided a recording of the tutorials, which related to the assessment piece. Links were provided in the marking guide.

7. Video of how to physically mark the assessment piece. As the marking is done online within the Learning Management System (LMS), a video of how to access the student’s work, use the rubric, and give audio feedback was provided for markers. A link for the video was provided in the marking guide.

8. Access to previous marking. Markers also have access to completed marking in previous semesters, so they could see additional examples of marker audio feedback. Links to specific examples were given in the marking guide.

Blind double marking early in the marking process (and randomly in the later stages). The blind double marking in this process occurred at three time points during the marking period by the two coordinators. To begin, blind double marking was undertaken when a marker had completed a small number of assignments (2-3 pieces). The marker received recorded audio feedback from the coordinator, regarding the accuracy of their grades and the quality of their feedback. If after 2-3 assignments have been double marked and accuracy and quality of the feedback were good, the marker was allowed to continue grading. However, if the coordinator was not satisfied, this process is continued (that is, a further 2-3 pieces are submitted to the coordinator for double marking) until the accuracy and quality of feedback were of an appropriate standard. Second, blind double marking was undertaken towards the middle of the marking period, and occurred when the marker has completed approximately half of their marking. This ensures the standard of marking is consistently maintained. Graded assignments were then chosen at random, and the same double marking process that occurred initially was commenced again. This included continued moderation if there are any issues. Also during this time, in addition to the random blind double marking, any fail grades are also double marked. Double marking fails as they occur, as opposed to at the end of the marking period, helps to ensure timely turn-around of assignments for the students. Lastly, at the end of the marking period, any fail grades that were (for whatever reason) not marked earlier, are double marked as per University assessment policy.

On-going development of marker expertise. A crucial component of the blind double marking is to provide markers with professional development. This personal communication occurs via the use of recorded audio files from the coordinator to the marker whenever double marking occurs. These audio files are used for developing marker understanding, skills, and ability to provide quality feedback in marking student assessment.

Typically, the recorded audio feedback files sent to markers: (1) focused on what the marker had done well, (2) discussed potential disagreements in grade or feedback for a criteria, (3) gave
examples of what was missed, and (4) offered examples of how feedback could have been enhanced. The marker then had an opportunity to discuss their understanding of the marking criteria via email or telephone, and was able to re-mark the student’s work. As mentioned above, the process was repeated (within the same marked assessment piece and / or with a different marked assessment piece) until both the subject coordinator and marker were confident that the marker was attaining the quality standard and consistency of all other markers.

As part of this feedback loop, markers were given regular updates (every couple of days) of the grade distributions of groups (approximately 65), the number of assessments marked for each group (and for the cohort overall), and the entire subject average. The purpose of this step is to create a feedback loop to enable the markers to judge how their marking compares to others.

Resources used for implementation
Student work was submitted and marked using the University's LMS, Brightspace by D2L. The LMS also provided information regarding marking averages that were disseminated to markers. Audio feedback to markers was recorded using the inbuilt computer device, QuickTime Player version 10.3 and emailed to markers via institutional email. Resource videos were created using Camtasia Version 2.10.1 and uploaded to YouTube for dissemination to tutors. YouTube videos were set to private and were only available to marking staff. Resources were shared with markers in a private space on the LMS. A team of two subject coordinators implemented the moderation practices.

Student and staff satisfaction
All 20 sessional markers were invited to participate in a focus group at the end of the semester. As many of the markers were geographically dispersed, only five agreed to participate. A research assistant, not working on the subject, conducted the focus group. Student satisfaction was surveyed at the end of the Semester via the University's internal student evaluations system.

Results
What did markers think of the moderation process?
In 2014 five markers (25%) were interviewed about the moderation process. Below is a summary of the discussion after the facilitator gave prompting questions.

Markers were asked how they found the process of moderation. All agreed that the processes were useful / helpful, particularly in helping them understand how well they were going. For example: “...I thought it was pretty helpful to know that you're either on the right track marking or if there was a question that you were a bit off on” (Marker Five).

Markers also agreed the process helped them understanding the marking criteria more clearly. For example: “It definitely helped me….I didn’t know how to differentiate [a criteria in the rubric], it helped clarify that for me going forward. So I did find that useful” (Marker Two).

Resources such as the marking guide, marking averages and marking meetings, were raised during the focus group. In particular, the marking guide and having the grade distribution sent around were seen as being very useful, although some found the group meeting recordings “pointless … because it was too many people talking at once …” (Marker Two).

Some other negatives raised included one marker feeling like (at least at the start of the moderation process) you could pick your best marking to be moderated. Others in the group pointed out that this was impossible with the random moderating practice that occurred later on, and it was “good that they’re keeping track” (Marker Three). One marker commented that they believed moderator gave feedback to them for the sake of it.

Markers were then asked if audio feedback from their moderator helped them to keep their marks calibrated and all agreed that it did: “Yes. It did, it was just really clear and it could point out which parts you needed to improve on, but also give you examples on how to say it better in your own feedback, so, it was good feed-forward feedback” (Marker Five). An interesting discussion arose where one marker made a comment regarding the link between the feedback they were receiving and the feedback they were giving to the students: “We do appreciate [the feedback]…. And so
do students so I’m thinking in that respect” (Marker Four). This viewpoint was shared by other markers in the room.

A limitation pointed out by one of the markers was getting feedback midway that didn’t really match their progress with marking: “I know it’s difficult because everyone marks in different rates but it could be more timely in accordance with the rate you are marking them to” (Marker Two). While other markers acknowledged that when they were running behind in their marking: “moderators were very forgiving with getting the feedback to you when you needed it” (Marker Four).

Markers were then asked whether having audio feedback from a moderator helped them to provide better feedback to the students. All agreed that they felt fairly confident with their own feedback, but felt it “it would be helpful for a newer marker” (Marker Five) and that it was more useful for “marking accurately” (Marker Four) and to “be consistent with the rubric because that ultimately is the most important thing” (Marker Two).

Markers were then asked if they had any suggestions for improving the moderation practice and the markers discussed the possibility of all marking the same piece of work and sharing their feedback. Suggestions included marking the same piece of work and listening to each other’s feedback, or going through the example as a group in a meeting. Even thought the markers recognised they could actually listen to the feedback of others already, only two said that they did this. However, they did acknowledge that new markers may be apprehensive regarding this prospect “perhaps we all need to get over ourselves and listen to each other’s verbal feedback” (Marker Two). They also suggested having a more informal meeting mid-way through the marking period to replicate the informal chats they were already having with other markers.

Were students satisfied with the feedback they received?
The University surveys students at the end of Semester on a number of satisfaction indicators. The below table shows (1) the satisfaction scores in 2010 for both large Health Psychology subject and the comparison University average at the time (Note: moderation process was not being implemented); (2) the satisfaction scores in 2011-2013 for the large Health Psychology subject in which the moderation practices were being implemented and refined; and (3) the current satisfaction scores in 2014 for both and the University average and the large Health Psychology subject in which this article is based. Note the increase in feedback satisfaction between 2010 and 2014; particularly the large increase in the number of students who strongly agreed that the feedback they received was helpful.

INSERT FIGURE 1 HERE

Discussion
In this large class context, this moderation practice was successful at enhancing both the accuracy, and quality, of the feedback given by markers. Particularly novel was the use of video and audio recordings to aid in moderation and to give formative feedback to markers. This feedback assisted marking staff to understand and accurately apply the marking criteria, as well as develops their skills in giving high quality, meaningful feedback to students. In particular, this moderation practice focused on: (1) good communication and resources with and for markers; (2) blind double marking early in the marking process (and randomly in the later stages); and (3) ongoing development of marker expertise. Each will be addressed in turn.

Good communication and resources with and for markers
The provision of resources facilitated a shared and accurate understanding of the marking criteria. Given the necessity of numerous sessional staff, and the geographically disparate location of staff and students, these resources were found to be particularly useful. As the marking guide contained most of the other resources, it was unsurprising that markers found it to be a “Godsend”, indicating a preference for having marking resources prior to marking, easily accessible, and located in one place. Within the marking guide, resources had varying use. The videos discussing marking criteria, the video of how to physically mark assignments online, as well the written feedback examples appeared more useful to new, less experience markers. Mostly, more experienced markers, who were already familiar with the marking and the marking process, did not use these resources. However, as less experienced markers developed confidence and their own style, they were less reliant on these scripts and videos.
From a moderator’s perspective, providing these resources appeared to help markers have a sound understanding of what represents ‘quality’ for any given assessment task. Marking meetings also provided an opportunity for markers to seek clarification, from the subject coordinators, on any issues to do with the marking process. However, the videoconferencing of these meetings was under-utilised by marking staff located on other campuses, and the focus group revealed the recordings were hard to hear at times. In future, it might be preferable to have some sort of post-recorded meetings, solely with the subject coordinators, going through the main points of the meetings.

**Blind double marking early in the marking process (and randomly in the later stages).**

The process of blind double marking was implemented to ensure markers were grading to the same standard. Starting moderation early, at the beginning of a marking period, has been argued by Bloxham et al. (2015a) to bring “greater consistency in advice, and encourage an ‘assessment for learning conversation’ amongst staff” (p13). In the present example, early use of blind double marking helped to shape conversations regarding any faulty beliefs in applying the assessment criteria. Blind double marking at the mid-point successfully detected differences in marker performance over time that may have been brought on by fatigue, tight timelines, or inexperience (Bloxham, 2009). Of added benefit was that double marking of fails was not left until the end of marking period, whereby it would otherwise slow the return of feedback to students (Jessop et al., 2012; Partington, 1994).

While blind marking at multiple time points improved the quality marking, it was challenging for subject coordinators to blind double mark such a sizeable cohort of sessional markers. It required both markers and coordinators to work on a schedule, so that timely feedback could be received. For example, markers were required to grade two assignments within the first 48hrs of the marking period commencing, enabling time to be scheduled specifically for initial moderation.

**On-going development of marker expertise.**

Important to this moderation process, markers expertise was continuously developed via audio feedback from subject coordinators. The use of technology to give audio feedback enabled each marker to receive (over several marked assignments) a sizeable amount of formative feedback regarding their marking. This was more than could be feasibly provided by written comments in the same timeframe. Formative feedback, as defined by Shute (2008), is the “information communicated to the learner that is intended to modify his or her thinking or behaviour for the purpose of improving learning” (p154). The focus group highlighted that the formative feedback helped them consistently apply the marking criteria, and check their understanding of its application.

Providing markers with audio formative feedback overcomes the time and location constraints that arise from engaging in face-to-face, individualised discussions (Jonsson, 2012). Studies in methods of feedback to students have shown that compared to written feedback, audio feedback can be less time-consuming, but provide significantly more detail than written feedback (Chalmers et al. 2014; Nemec and Dintzer, 2016). Lunt and Curran (2010) estimate that 6-minutes of writing is equal to one minute of audio feedback. Audio feedback has also been shown to contain more positive and more directive feedback (Nemec and Dintzer, 2016), is more personal (Neilsen, 2016), and often the preferred method over written feedback (Nemec and Dintzer, 2016; Neilsen, 2016). Audio feedback also allows for greater expression, tone, and nuance in the feedback (Carruthers et al. 2014).

The use of technology created efficiency of time, provided detailed responses, enhanced subject coordinator and marker interaction, and maintained a focus on marker development. This method improves ‘community building’ within the moderation process, where teachers share their understanding of the criteria and how they are applied (Bloxham et al., 2015a). Community building also has the added benefit of enhancing markers’ skills over the long-term, with less time needed for development in later assignments, and in future semesters.

Encouragingly, no post-hoc adjustment of grades for any one marker was required at the end of each marking period. This likely means the moderating process was successful at attenuating ‘hard’ or ‘soft’ markers early in the marking period. Audio feedback given by the subject
coordinator aided markers in modifying their marking practices, and enhanced the consistency, quality and accuracy of the markers. Obviously, from a quality standpoint this is important, but it also provides efficiencies to the marking process as well. Post-hoc adjustments can take a great deal of a moderator’s time if the moderator has to re-mark some or all of a marker’s graded assignments, more so if a marker has done a poor job grading a large amount of student assignments (in this scenario, upward of 100 individual assignments). Reducing post-hoc adjustments also means that feedback turnaround times are reduced, and students are able to receive feedback in a more timely manner.

While not a planned measure in this article, a surprising outcome was the paucity of student requests for re-marking, after grades had been released. Grade appeals typically happen when students are dissatisfied with their assigned grade (Germaine et al., 2011). Thus, the number of ‘change-in-grade’ requests can be used as an indicator of the quality of marking and feedback on assessment tasks. While a record was kept of the number of change-in-grade requests during 2014, comparison records prior to this time were not kept. Further, a search of the literature did not find any data that could be used to compare the frequency of change-in-grade requests. That being said, in total, students made only eighteen change-in-grade requests from 1500+ students across three assessment pieces. It is reasonable to surmise that the small number of requests speaks to the quality and transparent integrity of the process for both markers and students.

The small number of change-in-grade requests is complemented by the level of satisfaction students had with the feedback they received. Typically, in the university sector, students are dissatisfied with the quality of the feedback they receive (Yorke et al., 2010). However, for the Health Psychology subject illustrated in this study, student evaluation scores indicate that compared to pre-moderation practices (2010) or University average (2014), the 2014 Health Psychology student cohort was very satisfied with the feedback they received. In particular, more students indicated that they strongly agreed that the feedback was helpful. It appears from first blush that the moderation process improved the consistency and quality of feedback, and this is reflected in satisfaction scores given by students. This is a positive outcome for both staff and students alike, and is encouraging for moderation practices for large volumes of assessment.

The most obvious limitation of the moderation process is the time commitment for members of the teaching team. This time commitment includes the time required setting up all the resources, and the time to give audio feedback multiple times for each marker. Setting up the resources initially is time intensive. It is also acknowledged that should any assessment task change, the video and/or audio recording(s) might need to be re-done, thus adding to the workload. However, most resources can be reused with only small modification, if set up correctly. For example, in reference to the marking criteria videos, if parts of the marking criteria change, that section of the video is removed, and only that section replaced. While it can be time-consuming, it is easy to argue the initial time outlay is worth ensuring that markers clearly understand the task at hand.

Giving audio feedback multiple times can also be a time-consuming process for the subject coordinator. However, developing the marker’s skills early, means markers quickly become self-sufficient and make fewer mistakes. As marking staff became more skilled, they required less feedback, and there was a noticeable increase in the efficiency of the moderation process. In addition, increasing marker expertise resulted in time saved not implementing post-hoc adjustments. This has benefit not only to the current semester’s marking, but also future assessments, where markers enhanced marking skills can be used again, with less time needed for development.

There also is the possibility that the use of marking distributions increased the likelihood of markers being risk adverse and erring towards the cohort average. The mid-point marking did help to alleviate this challenge somewhat, as it reinforced high or low grading if it reflected the group performance, rather than a deviation by the marker. This helped to discourage markers erring towards the cohort average, if it was not a true reflection of the performance.

Lastly it must be noted that the student evaluation question regarding feedback did change in 2014. The question asked to students in 2010-2013 was ‘The teaching staff gave me helpful feedback’, which was changed to ‘Feedback on my work in this subject helps me to achieve the learning outcomes’ in 2014. While the question is slightly different, and may have affected how
students responded, the student satisfaction scores presented during the period of 2011-2013 show that while the moderation process was being refined, there was still a large increase in student satisfaction in comparison to 2010 (prior to implementation).

Markers did make suggestions to improve this process. They were particularly interested in trying group moderation or consensus marking, where they were able to all mark the same piece of work and share feedback. To circumvent the issue of markers being geographically spread out, the use of a written practice example would allow markers to apply the criteria to the same piece of work (real previous assignments), and check their answers and understanding against the feedback the example student received. There is certainly scope in the future to build upon this resource and perhaps find ways to anonymously share feedback between markers.

While more research is needed to balance workload challenges, overall, this article highlights the importance of a shared understanding of marking criteria for quality assurance. This is particularly important in an Australian context where the majority of teaching academics are casually employed, with less opportunity for professional development (Crimmins et al., 2015). Future research should focus on the implementation of audio feedback for community building, as well as enhancing shared understanding. While an important practice, Bloxham et al. (2015b) argue more evidence is needed about the sustainability and impact of these practices.

In large class teaching in particular, systematic methods of moderation need to be implemented. These methods should include: (1) resources / communication that are provided to markers in order to facilitate a shared understanding and interpretation of the marking criteria; (2) multiple points of double marking to detected differences in marker performance over time that may have been brought on by fatigue, tight timelines, or inexperienced; and (3) develop marker’s skills early through formative feedback to acquire self-sufficiency, accuracy, and expertise in the in the grading process. The practice of moderation utilised in this article overcame some of the challenges of moderating large volumes of assignments. Particularly, the use of audio feedback, and video resources were the most novel and useful. Undeniably this article is limited in evidence, and is just a starting point for discussion about how to increased quality, efficiency, and equity in the assessment process while maintaining a focus on marker development, in large class moderation practices.

References


Figure 1. Student evaluation scores regarding whether the feedback they received was helpful.