Enhancing assessment literacies through development of quality rubrics using a Triad based peer review process

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

Enhancing assessment literacies through development of quality rubrics using a Triad based peer review process

Abstract

Peer review is viewed as a valid quality assurance mechanism in higher education. Peer review of teaching is common practice at universities in Australia. However, peer review of assessment is a relatively new innovation in tertiary education. Peer review of assessment in a Triad structure utilised data, via interviews with academics and students, to develop a peer review of assessment framework. This project was modelled on a Triad based peer review of teaching process at a major university in Brisbane Australia. A 10 question framework was used initially to facilitate conversations between assessors in a range of undergraduate courses (teacher education, business, visual arts, occupational therapy, outdoor education). The benefits for all stakeholders were widespread and significant, impacting students and assessors and provided a response to sector wide, national and international criticisms of tertiary assessment by students, who are driven by successful assessment experiences and shaped by the clarity of assessment rubrics.

Keywords

Introduction

Peer review is increasingly being viewed as a valid quality assurance mechanism in a new era of accountability in higher education. Peer review of teaching is now common practice at many universities in Australia (Booth et al, 2016). However, peer review of assessment is a relatively new innovation in tertiary education and driven by student dissatisfaction with many assessment practices. This accountability is characterised by an emphasis on a quality tertiary education experience and partly influenced by official student evaluations of courses and programs. Students criticise assessment in higher education as being subjective, norm-referenced, characterised by academics mystifying assessment processes and criteria, and confirming a negative image of, even negligent, assessment practices (Ecclestone, 2001).

This increasing emphasis on quality is policed by the tertiary 'watchdog' in Australia known as Tertiary Education Quality and Standards Authority (TEQSA). The most striking example of this is a forthcoming requirement for tertiary education providers to benchmark similar course and programs between institutions. A sustained focus on assessment quality is a natural consequence of the necessity to quality assure courses and programs and this is being secured by a focus specifically on assessment practices and the artefacts of assessment, including grading tools also known as criteria sheets or Guides to Making Judgments (GTMJ). It is well known that student engagement is driven by successful assessment experiences (Boud and Associates, 2010) and shaped by the clarity of assessment rubrics (Grainger & Weir, 2016). In an attempt to provide a response to these sector wide, national and international criticisms of tertiary assessment, researchers have, to date, generally focused on investigating and reporting on the process of providing assessment feedback, given by assessors (Carless, 2006) rather than the assessment rubric itself (Grainger, 2015b).

Literature Review

Rubrics typically contain three elements: criteria, standards and standards descriptors. Standards descriptors are supposed to be precise verbal descriptions, allowing for unambiguous determinations, consisting of statements setting down the properties that characterise something of the designated levels of quality (Sadler, 2009). The description of quality identified in each of the standards is matched against evidence in student work. 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. Hence, grading tool deficiencies represent major challenges to what Sadler (2010) refers to as grade integrity.

A rubric is not only used to determine the standard of student work, it also acts as a feedback mechanism because it links the evidence in student work to the standard descriptors on the grading tool, hence providing students with guidance as to why a standard was received. 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 & William, 1998; Ramsden, 2003). However, although there have been numerous studies carried out into the effectiveness of feedback for improving student learning, empirical evidence shows that while feedback communications may be appreciated by students, they often lead to little if any improvement in subsequent submissions (Sadler, 2009). As a consequence, feedback research in recent years has focussed on the student as an active agent in the feedback process (Carless & Boud, 2018; Carless, 2019a, 2019b; Henderson et al, 2019; Dawson et al, 2019; Malecka et al, 2020; Mercader et al, 2020; Carless & Winstone, 2020; Winstone et al, 2020; Winstone & Boud, 2020).

A further key assessment process is moderation, also impacted by the quality of the rubric. Moderation occurs post grading when assessors compare assessment judgments about the same

student work for purposes of calibration, which ensures consistency of judgments. 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). All these significant assessment related issues involve the use of rubrics.

Therefore, the rubric must be viewed as the key mechanism for the quality assurance of assessment, and therefore, fundamental to improving all stakeholders' assessment experiences and ultimately, student outcomes. Assessors and students are impacted prior to grading as they engage with the rubric and deconstruct its meanings. Assessors are impacted at the point of grading as the rubric is used to match evidence of student work with standards and standards descriptors. Assessors are also impacted during the moderation process which uses the rubric to ensure consistency of judgments, and finally, students are impacted post grading when they receive results and guidance in the form of feedback via the rubric. If the rubric is vague, students are reliant only on assessors' annotations for guidance, and these too, are the subject of severe criticism as being too difficult to decode (Sadler, 2009). The rubric is, therefore, the most significant assessment artefact and yet, it has not been a major focus for research in the tertiary sector (Grainger & Weir, 2020).

Research on rubrics, as noted above, is scarce in the tertiary sector, or specifically, research comparing models of criteria sheets is rare. Rubrics are used to decide grades. Deciding on grades is one aspect that is particularly problematic (Sadler, 2005). Universities in Australia have various procedures for awarding and monitoring the distribution of grades to students but criterion referencing is now common in most if not all tertiary institutions in Australia. Criterion referencing attempts to establish grading decisions by reference to some external and relatively fixed set of criteria. In principle, a student's work could be judged against the relevant criteria in isolation from the work of other students. In addition to criteria, a focus on standards is now commonplace giving rise to what is commonly referred to as a criteria and standards based assessment model, characterised by rubrics that include criteria (what is valued in the assessment task) and standards (typically High Distinction, Distinction, Credit, Pass and Fail) and descriptors of achievement for each of these standards (standards descriptors).

Despite this focus on criteria and standards, assessment of student work remains a relative mystery to both students and assessors alike (Haines, 2004). One of the major reasons for this is a failure to create a quality rubric characterised by clearly identifying criteria for achievement and explicit descriptors of standards of achievement for each of the criteria. Whilst criteria are generally quite specific in identifying aspects of behaviour about which to make assessment judgements, typically, descriptions of standards are often subjective and open to interpretation. Often the standards descriptors themselves provide insufficient guidance for markers (or assessors) for those 'close calls' which are often differentiated by the use of subjective terminology like 'extensive', 'significant', 'comprehensive', 'outstanding', 'adequate', 'expert', 'insightful', 'discerning' and 'good'. The key difficulty is often that the real standards are locked inside the marker's head as tacit knowledge and have not been explicitly stated to students. This is not only a problem for students when interpreting the expectations, but a problem for assessors when moderating. Standards descriptors that fail to discriminate standards are a significant problem with many rubrics, resulting in inconsistency of assessment judgments by academics.

As a result of these deficiencies, and despite the existence of common criteria and known standards, identified in rubrics, markers mark the same piece of assessment differently (Grainger, Purnell, & Zipf, 2008). In assessing the quality of a student's work or achievement, the assessor must possess a concept of quality appropriate to the task, and be able to judge the student's work in relation to

that concept (Popham, 2005). In other words, a strong understanding of the knowledge, values and expectations is needed to ensure the assessor's judgment is also of quality. This knowledge can only be utilised if the assessor has had the experience of marking the task previously, in short, has the tacit knowledge (Sadler, 2010).

Academics who are experienced assessors possess tacit knowledge of what quality looks like in student work because competent appraisers can consistently identify quality when they see it (Sadler, 2011). This tacit knowledge has been shown to enable assessors to make accurate interpretations of sometimes vague descriptions of student behaviour in order to discriminate between standards or levels of achievement (Grainger, Purnell, & Zipf, 2008), which means there is no compromise of assessment integrity and reliability in terms of assessor judgements. However, not all academics understand or are experienced with sound assessment practices and it takes some years to get to know how to align evidence of quality with relevant achievement standards and achieve consistency of judgement over time. For assessors who are unclear about learning quality, vague assessment grading tools are not, in fact, objective arbiters of performance, nor are they defensible, nor do they encourage consistency of teacher judgements.

In this way, even an experienced assessor, who marks a task for the first time, can be referred to as a novice assessor (Grainger & Adie, 2014). Where there are multiple markers, and even multiple novice markers, the results of marking and moderation are even more problematic. Hand in hand with this, and prior to grading, is the very real issue of creating a new rubric for an assessment task, for the very first time. That is, without experience, the rubric creator can only imagine what behaviours will be characteristic of each standard for each criterion. These behaviours are reflected on and modified over time as a result of multiple marking experiences, eventuating in a quality rubric that explicitly describes standards for each of the predetermined criterion. This process takes time.

The rubric has multiple roles. Primarily, grading tools are designed by and for assessors to evaluate the quality of student work. They also enable assessors to construct feedback to students that aligns with the criteria for marking. From a pedagogical perspective, the grading tool can act as a guide for students' learning by making explicit the evidence they must demonstrate through the assessment task. Because grading tools have more than one purpose and audience they are complicated texts to construct, especially the standards descriptors to ensure they adequately differentiate between the levels of achievement. This often results in standards descriptors that are ambiguous, open to interpretation and make assumptions that the user is familiar with the language used (Sadler, 1987).

A literature review failed to reveal a similar project in the tertiary sector. A literature review revealed only one study that utilised a peer review structure (Grainger, Bridgstock, Houston & Drew, 2015), but this peer review was based on observations of teaching in a triad structure, and not assessment artefacts such as a rubric. The reported peer review of teaching structure had operated at the university for over 10 years and is based on the premise that a quality evaluation of teaching is best achieved via a triad structure consisting of a peer, a discipline expert and the reviewee. That structure has been duplicated in this project and is the rationale for implementing it.

This project investigated the development of assessment literacy, specifically the ability to create quality assessment rubrics, in teaching academics across a range of disciplines (teacher education, business, visual arts, occupational therapy, outdoor education) over a two year period. The project utilised a triad structure, consisting of eight Triad Leaders (experts in assessment from the discipline of Education), eight course coordinators and eight tutors teaching the same courses from each

discipline. The Triad conversations were based on a 10 question framework. The key research questions were:

- What are the strengths of a Triad based structure for peer review of assessment?
- What are the strengths of the 10 question framework for peer review of assessment?

Methods

The project utilised a qualitative research design, namely an Action Research Approach (Bryman & Bell, 2011). Specifically, the approach used a reflective process of problem solving, led by the Project Leader, working with seven other academics in Triads, as part of a "community of practice" (Wenger, 1998) to improve assessment rubrics. This approach was an extension of a similar assessment project led by the author to develop academic staff assessment literacies through assessment focussed conversations. This approach was chosen due to its cyclic, participative and reflective nature. Discussion/interview was the principal data gathering method. The conceptual framework for the study was based on a sociocultural view of learning as mediated practice. Lave and Wenger (1989) identify the concept of 'situated learning' as learning that is embedded within contexts, culture and activity. A ten question framework (Christie et al 2015) was utilised by Triad leaders to initiate, guide and sustain conversations about the rubric with other Triad members.

Triad-based activities were undertaken over a two or three semester period. Participation was voluntary and determined by the level of assessment literacy possessed by the Triad leader, evidenced by ongoing participation in assessment projects led by the author at the School level and assessment focussed publications. The Triad included an 'expert' from the Education discipline and two self-identified academic peers who worked collaboratively as a mini community of practice, to improve assessment artefacts over the course of two semesters. Triad leaders did not know other members of their Triad. In total eight Triads were utilised in the project. 25 hours of taped transcripts were generated. Each Triad participated in two or three meetings over the duration of the two year project. A total of 11 students from five different triads were interviewed in the final phases by the Project Leader (author). The student interviews were transcribed resulting in 58 typed pages and seven hours of student responses. Finally, a one day workshop, involving all Triad leaders was held to discuss the results. Student results are not the focus of this paper and will be reported elsewhere (forthcoming). Ethical clearance was given by the university, Ethical Clearance Number A181065.

Data Collection Methods

There were multiple phases gathering multiple sets of data from Triad leaders, triad members and students:

- Evaluating an existing assessment task and rubric through a focus group structure involving the Triad;
- Rewriting the existing rubric to ensure constructive alignment and standards based on the Triad conversations;
- Trialing the improved rubric with students in selected courses
- Evaluating the improved rubric through follow up Triad based focus group discussion centred on how the improved rubric impacted assessment processes.
- Interviewing students to evaluate the effectiveness of the rubric
- Collecting data from Triad leaders via on line survey as a lead in to the final data gathering phase, a professional development workshop

- Collating and sharing experiences with each Triad in a one day professional development workshop, resulting in the honing of the existing framework and the development of a new framework for peer review of assessment.

Results

The results are presented based on three data sources: recorded Triad conversations; survey analysis of qualitative comments from Triad leaders and conversations between Triad leaders recorded during the professional development workshop at the conclusion of the project.

Results from recorded Triad conversations

Triad conversations were recorded, transcribed and analysed for themes by the author initially and then confirmed as accurate by Triad leaders independently.

Conversations with Triads led to many interesting, challenging and sometimes 'heated' exchanges as to how to construct rubrics, level of specificity, agency, ownership, impost of authority, resulting in deep reflection as to the art and science of creating quality rubrics. A major focus for discussion was the subjective language used in standards descriptors. On all occasions the revised rubric has been a much clearer and improved guideline for both markers and for students, encouraging an exchange of communication as to the requirements of assessment tasks that had not been evident in the original rubrics.

Staff members reacted differently to the conversations, some perceived them as challenges to existing work, others embraced the opportunity to develop new skills and improved assessment outcomes for students, as a result of developing a rubric with clarity. The Triad structure worked well, particularly with Triad members who had not known each other prior to the Project. This allowed sometimes 'difficult' messages to be conveyed by the Triad leader. The nested aspect of rubrics caused some issues to occur with some Triad members. Some Triads did not see the necessity of creating explicit rubrics because of the existence of detailed guidelines/instruction sheets/scaffolding and hence there was some 'pushback'. Having different disciplines enabled Triad leaders to develop an understanding of the specific needs of some disciplines that were not from the School of Education. For example, one Triad was based on Visual Arts and the Triad leader was from the discipline of Education.

In this regard (professional development) all noted great professional development experiences through substantive conversations and through engagement with academics from other disciplines. It was the dialogic nature of the discussion that enabled the professional development opportunity for all involved in the process. Many noted the difficulty of writing good rubrics and the need for allocating time to do so. They valued the opportunity to describe to each other what students did at particular levels and then these ideas were crafted into descriptors for each standard. The opportunity to express different ideas and perspectives were canvassed. By voicing concerns and questions all members of the Triad increased their understanding of assessment and the ways in which students can be assisted to understand what the task required and how the quality of their response will be judged.

Some identified the need to widely disseminate the Triad experiences and mentor other academics who had not participated in the project. In some instances, the positive outcomes were sustained as participants developed better rubrics subsequently, once the Triad was dissipated. This impact or influence was not consistent however, and it was agreed that rubric reviews should be part of an

ongoing cycle of development and redevelopment and probably connected to the review of course outline process that occurs annually.

All Triad members valued the open processes of discussion and sharing between academic staff, student and tutors and provided a context for ongoing refinement and development of assessment literacy. All Triad members valued the use of 10 question framework which proved to be a very relevant framework for initiating conversations about rubrics with participants. The main key to quality assessment rubrics was identified as vocabulary: the use of explicit and behavioural vocabulary that indicates what is truly valued in the assessment task; consistency with the use of words; differentiating the standard descriptors at different levels; verb selection for standards descriptors is important in creating explicit descriptors; subjective versus explicit wording in descriptors – a conversation about 'semantics'; vagueness of rubrics; inconsistency of assessor judgments; the necessity to align the words used in the assessment artefacts (rubric, guidelines) with the published course outline.

The course outline was a constant theme for discussion, or more specifically, alignment to the course outline. It was noted that course outlines were written well in advance of rubrics sometimes as long as 12 months. The alignment between the course content, learning outcomes, task requirements and criteria were challenging parts of the conversation to ensure everyone had a vision of how these connected.

Conversations about what to expect were easier to explore on tasks that academics had undertaken and marked previously. This would be very challenging for new tasks as evidenced in the rubric for a newly developed task in this course.

All triad members noted common purposes of good rubrics as supporting student learning and success in the task; providing provide usable feedback; shaping the task for the student especially the characteristic of nestedness. All agreed that it was significant to discuss rubrics with students as well as the tutors teaching the course and marking using the rubric. All agreed good rubrics were useful when justifying or defending marking decisions

Results from survey analysis

As noted previously, Survey Monkey was used to gather responses to a series of questions prior to the implementation of the professional development workshop, the final phase of the project. Representative responses are identified and presented to illustrate key messages in relation to survey questions. Analysis of the responses indicated a triangulation of results and a remarkable consistency of key messages, with some differences.

The key themes were: improved clarity for markers and students due to descriptors that clearly outlined the evidence needed to achieve a certain standard; better alignment between assessment task and course content as the rubric guided the delivery of content for tutors; positive impact on moderation processes as markers were on the same page due to the clarity of behaviours identified in the descriptors; increased clarity as rubrics became more succinct with less noise; increased consistency of teacher judgement and accessibility of students to the task resulting in improved student satisfaction emanating from fewer disputes and improved results; the development of a shared assessment metalanguage between assessors leading to greater transparency of the task and criteria sheet and the overall development of assessment literacies.

Results from the professional development workshop conversations

A professional development oriented one day workshop was held at the end of the project. All Triad Leaders were invited. The workshop aimed to summarise the findings of the project and to synthesise the achievements to accomplish the final objective of the project, the establishment of a revised framework to guide the development of assessment rubrics. The conversations were recorded, and a Research Assistant was employed to transcribe the conversations. Once again, word clouds were generated by Survey Monkey software to confirm the identification of nine themes by the Project Leader. The themes identified, without intention of privileging them, were: Relationships; Tacit knowledge; Academic Agency; Discipline specific issues; Stifling of creativity; Ongoing process of refinement; Expertise in schools on rubric development; The utility of a framework to reveal strengths and deficiencies in rubrics; and The future of rubric resources. The relevance of these themes are illustrated in the synthesis of results provided below.

Many of the issues that were revealed in regard to the construction of explicit rubrics can be subsumed under the theme of relationships. Shared thinking of Triad teams (i.e. course coordinator and tutors) was a hindrance in that some teams knew what each other expected from students, but that wasn't made explicit in the rubric. A shared understanding comes about via the development of tacit knowledge over time. Some tutors hadn't realised that some words were ambiguous or subjective - for example, the word *insightful*, was not considered ambiguous nor subjective because they knew what they had intended when writing the descriptors and knew what to look for in student work. Whilst this shared understanding ensured a consistency of teacher judgments, this understanding was not shared by students, as revealed in all of the interviews with students conducted by the Project Leader. That proved to be a difficult and sometimes protracted process to work through, but once the tutor understood what it meant, she was able to move towards more explicit language in the rubric. It was challenging for the tutor to realise that only her knowing what was meant was not sufficient. The lesson is that language used in rubrics needs to be explicit - especially when dealing with large courses with many tutors. Deference to the Course Coordinator sometimes, but not always got in the way of the tutor giving clear feedback to the Course Coordinator in regards to the clarity of the rubric. Significant concepts here are power relationships, deference to authority and the tenuous position of casual staff and how tutors perceived their agency. Agency is developed over time (ie continuity) and can be difficult to sustain in times of rapid change as the university has experienced over the past 3 years at least. 'Inherited courses' as a result of staffing changes can cause problems with rubrics due to this lack of shared tacit knowledge between new teams. In summary, issues with continuity of staff hinders buy in and derails agency. These rapid changes impacted directly on two Triads, both of which were either unable to implement the revised rubric or were forced to start again as courses were discontinued, and as course coordinators were changed.

An advantage of the Triad structure was the fact that the Triad leader did not know other Triad members. In this project two of the Triad leaders were from outside of the tertiary environment, two practising school teachers with currency in assessment. This was certainly a success factor for those Triads as they brought fresh eyes and assessment currency. As an 'outsider' the Triad leader was in a position to have the critical lens and there was emotional safety for all participants. Some teams were egalitarian (eg OT), in fact, some tutors had more of a handle of the task requirements and hence the construction of the rubrics was collaborative, in some occasions the CC deferred to the tutor.

This project was driven by the need for clear rubrics. However, the need for clarity also stifles creativity by students who might perceive there is only one 'formula' identified in the rubric as a set

of behaviours that enables a certain standard to be achieved. A good example of this was the Visual Arts Triad. This Triad was very successful, but it did take time initially, as the Triad leader came from a Humanities/History/Languages discipline. Although there are advantages to being prescriptive there are also disadvantages, as it stifles creativity - especially in design – so the process for rubric development in this Triad had the added pressure of how to manage that tension between creativity and clarity. An explicit, supportive document can 'derail creativity', so it is a fine balancing act to create a rubric that is both explicit and yet provides enough flexibility to enable multiple behaviours for a standard to be achieved. In the case of business, the emphasis was on providing a rubric in addition to a set of explicit guidelines that teased out the rubric. The issue with this was that students did not always attend class to be part of the conversations so that places an emphasis on the rubric as being the main guide for students. This was reiterated especially in first year when students do not have the contextual knowledge to understand the rubric alone without unpacking it through dialogic conversations with tutors.

A theme that emerged from workshop conversations is sustainability. That is, are the lessons that have been learned as a result of being part of the Triad sustainable, and capable of being reproduced once the Triad disbands? In some cases, this proved to be a challenge as the second rubric created, without the explicit input of the Triad leader, proved to exhibit the same original negative characteristics and habits. In short, the process of rubrics' refinement must be ongoing. Academics can see writing rubrics as a one-off task, not a continuous cycle of improvement and there is a need to shift the mindset so that the rubric review is part of the annual course review process. New assessment tasks are challenging when creating clear rubrics due to a lack of tacit knowledge, new rubrics, created for the first time can only identify imagined behaviours because of the lack of tacit knowledge about what to expect from students at various standards.

The next major theme to emerge from the workshop conversations focussed on the future. Good quality rubrics need trained rubric makers and some Schools do not have a tradition of understanding how to write rubrics as they are content experts but not expert in pedagogy. (SoE does). Having a common framework is hence, a good strategy and there was general discussion that the existing framework was successful in this project because it revealed shortcomings, specifically, discrepancy between what people valued in the task, opened the marking up to inconsistency of judgement. Having a framework however was not enough to guarantee success. Significant in this discussion was the need for Triad leaders to be good facilitators, the conversation broker needs to be an expert, knowing how to create and lead Triad conversations, how to broker conversations that build relationships and trust, build commitment to improvement, develop momentum and then trigger 'agency'. If this does not occur, then it is the 'blind leading the blind'. In terms of sustainable improvement one Triad leader noted that we haven't created experts in this process, we've created people who are reliant on experts and a one off experience did not change practice. Multiple cycles are needed to effect sustainable change to the development of quality rubrics across the whole university. In this regard, there was considerable discussion on using existing expertise, possibly as a result of this project, to supplement and complement existing resources developed by CSALT, possibly linking to incentives, such as promotion or HEA fellowship pathways.

Discussion

As noted in the introduction to this paper, a review of the literature revealed only one study that utilised a peer review Triad based structure, and that process focussed only on live teaching, not on assessment. Hence, there is no literature which can be used to evaluate and the results of this study. Therefore, the discussion of results of this study are framed in terms of answers to the key research questions (identified earlier) that will provide guidance for commencing the process of peer review

of assessment using a triad based structure. The narrative is based on a synthesis of data gathered from all the data gathering mechanisms of this project, namely, Triad conversations, survey data and the final professional development workshop conducted at the end of the project. It should be noted that this kind of peer review process is an internal process within institutions and cannot be equated to benchmarking externally.

A major strength of the Triad based peer review process is that it allowed course coordinators, independent assessment experts (Triad leaders from another discipline), tutors and students to work collaboratively to make positive improvements to the assessment task and accompanying rubric as well as strengthening alignment to the teaching program to support students. This is based on the fact that that the members of the Triad have independent knowledge, skills and perspectives on the development of rubrics and hence they bring to the focus group table fresh eyes, a plethora of experience and insights that would not normally be probably due to existing structures that preclude such collaboration.

Triad structures are not guaranteed of success. Successful Triads in this project were characterised by respect for each other and for the roles that were being undertaking. Successful triads were characterised as democratic and the conversations flowed freely. Successful Triads focussed on the development of academic agency (Heck, 2019) or ownership, and the skill of the Triad leader as a facilitator of this emerging agency. The key to this development rests on a trusting collegial relationship between the Triad members, and a willingness to adopt new practices. This trust can be established easily if the Triad leader is hitherto unknown to other Triad members, facilitating the delivery of messages that may have been undeliverable due to existing relationships between teaching teams, hindered by imposts of authority. Time to develop these relationships is of utmost significance and time to implement changes is of equal significance. The development of academic literacy is time consuming, and a cycle of is required to ensure any positive changes to practice are sustained and sustainable. This may involve allocations to workloads to ensure buy in and avoid buy out and a return to familiar practices, in short adopting a gradual release model that develops agency.

In relation to the 10 question framework, as a result of the data presented above, some questions were retained and reworded, others were deleted and finally, it was decided that a series of questions before the Triad meetings, during the Triad meetings and after the Triad meetings would be a better framework/protocol. The following revised protocol is suggested, reflecting these three stages.

Before the Triad: Reflect on a previous task and rubric from the perspective of the course coordinator, tutors, marker, moderator and student. From each of these perspectives:

- What worked well in the task and the rubric?
- What were concerns regarding the task and the rubric?
- How good was the rubric when the student work was graded, moderated and when feedback was given?
- How do you know that your markers have a common understanding of how to mark according to the rubric?
- What are the existing processes of review you have in place for refining rubrics?

During the Triad meeting: Share perspectives on the task and rubric with the Triad.

1. What information is given to students in lectures or tutorials to support their understanding of the task and rubric?

2. Does the task provide opportunities for the students to demonstrate that they have achieved its intended outcomes?
3. Does the rubric align to the course outline?
4. Does the rubric contain criteria, standards (e.g. HD, D, C, P, F) and standard descriptors?
5. Are the criteria and standard descriptors unambiguous?
6. Do the standard descriptors include verbs that clearly articulate what is required to meet the standard?
7. Are the standard descriptors explicit and devoid of subjective language?
8. Are the standard descriptors positively worded in terms of what students must do?
9. Are there gradations of quality that differentiate the standards clearly according to a recognized taxonomy of learning?
10. Are aspects of the standard descriptors repeated? Would the rubric benefit from becoming a nested model?

After the Triad process has been completed make necessary changes to the rubric and share this with staff. Review the updated rubric following the same Triad process.

After the Triad

- Implement the task and rubric
- Create support materials for students and staff to facilitate understanding of the task and rubric
- Create a process to gather feedback from the course coordinator, tutors, marker, moderator and student on the effectiveness of the rubric.
- Repeat the Triad process

Conclusion

This Triad structured project was modelled on a very successful Triad based peer review of teaching mechanism operating at Griffith University in Brisbane (Grainger, Bridgstock, Houston & Drew, 2015). A 10 question framework (Christie et al, 2015) was used initially to facilitate conversations between assessors in range of undergraduate courses (teacher education, business, visual arts, occupational therapy, outdoor education). A focus on higher education professional learning experiences for academics, through a Triad structure, on a subject that is an important aspect of the student experience, is significant. The overarching outcome of this project is changes to the learning and teaching experience in the form of a greater understanding by assessors of key assessment processes, namely, crafting assessment tasks and criteria sheets for grading, moderation and feedback. This was accomplished by a focus on changing practice and using reflective practice as an approach to achieve objectives, enable professional learning opportunities, better student outcomes, improved accountability evidence, and greater staff and student satisfaction due to improved assessment experiences. Utilising a peer review process to evaluate the quality of rubrics has led to enhanced stakeholder understandings of assessment due to the engagement required and the development of 'agency', thereby building capacity and ensuring sustainability.

References

Booth, S., Beckett, J. & Saunders, C. (2016). Peer review of assessment network: supporting comparability of standards. *Quality Assurance in Education,* 24(2), 194-210. https://doi.org/10.1108/QAE-01-2015-0003

Boud, D. and Associates (2010). *Assessment 2020: Seven propositions for assessment reform in higher education*. Sydney: Australian Learning and Teaching Council. Available at https://www.uts.edu.au/sites/default/files/Assessment-2020_propositions_final.pdf

Bryman, A. & Bell, E. (2011) *Business Research Methods.* 3rd edition, Oxford University Press, Oxford.

Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education*. 31(2), 219–233. https://doi.org/10.1080/03075070600572132

Carless, D. & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315-1325. https://doi.org/10.1080/02602938.2018.1463354

Carless, D. & Winstone, N. 2020. Teacher feedback literacy and its interplay with student feedback literacy, *Teaching in Higher Education*. https://doi.org/10.1080/13562517.2020.1782372

Carless, D. (2019). Learners' Feedback Literacy and the Longer Term: Developing Capacity for Impact. In: Henderson M., Ajjawi R., Boud D., Molloy E. (eds) *The Impact of Feedback in Higher Education*. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-25112-3

Carless, D. (2019). Feedback loops and the longer-term: towards feedback spirals, *Assessment & Evaluation in Higher Education*, 44(5), 705-714, https://doi.org/10.1080/02602938.2018.1531108

Christie, M., Grainger, P., Dahlgren, P., Call, K., Heck, D., & Simon, S. (2015). Improving the quality of teaching and learning in Master of Education courses by modelling best practice in the writing of assessment criteria sheets. *Journal of the Scholarship of Teaching and Learning,* 15(5), 22-35. https://doi.org/10.14434/josotl.v15i5.13783

Dawson, P., Henderson, M., Mahoney, P., Phillips, M., Ryan, T., Boud, D. & Molloy, E. (2019). What makes for effective feedback: staff and student perspectives. *Assessment & Evaluation in Higher Education*, 44(1), 25-36. https://doi.org/10.1080/02602938.2018.1467877

Ecclestone, K. (2001). 'I know a 2:1 when I see it': Understanding criteria for degree classifications in franchised university programmes. *Journal of Further and Higher Education,* 25(3), 301-313. https://doi.org/10.1080/03098770126527

Grainger, P. Purnell, K. & Zipf, R. (2008). Judging quality through substantive conversations between markers. *Assessment & Evaluation in Higher Education,* 33(2), 133-142. https://doi.org/10.1080/02602930601125681

Grainger, P. and Weir, K. (Eds). (2020). *Assessing learning in higher education contexts using rubrics*. Cambridge Scholars Publishing. Newcastle upon Tyne, UK.

Grainger, P. (2015). How do pre service teacher education students respond to assessment feedback? *Assessment & Evaluation in Higher Education*, 45(7), 913-925. https://doi.org/10.1080/02602938.2015.1096322

Grainger, P. &Weir, K. (2016). An alternative grading tool for enhancing assessment practise and quality assurance in higher education. *Innovations in Education & Teaching International*, 53(1), 73-83. https://doi.org/10.1080/14703297.2015.1022200

Grainger, P. & Adie, L. (2014). How do pre service teacher education students move from novice to expert assessors? *Australian Journal of Teacher Education*, 39(7). https://doi.org/10.14221/ajte.2014v39n7.9

Grainger, P., Christie, M., Thomas, G., Dole, S., Heck, D., Marshman, M. & Carey, M. (2017). Improving the quality of assessment by using a community of practice to explore the optimal construction of assessment rubrics. *Reflective Practice*, 18(3), 410-422, https://doi.org/10.1080/14623943.2017.1295931.

Grainger, P. Heck, D. & Christie, M. (2017). Evaluating the efficacy of Masters coursework criteria sheets using a 10 question framework. *Journal of Further and Higher Education*, 43(3), 396-405. https://doi.org/10.1080/0309877X.2017.1363384

Grainger, P., Bridgstock, M., Houston, T. & Drew, S. (2015). Working in Triads: A case study of a peer review process. *Journal of University Teaching and Learning Practice,* 12(1). http://ro.uow.edu.au/jutlp/vol12/iss1/3

Haines, C. (2004). *Assessing Students' Written Works: Marking Essays and Reports*. London, England: Routledge.

Hattie, J., Biggs, J. & Purdie, N. (1996). Effects of Learning Skills Interventions on Student Learning: A Meta-Analysis. *Review of Educational Research*, 66(2), 99-136. https://doi.org/10.3102/00346543066002099

Henderson, M., Phillips, M., Ryan, T., Boud, D., Dawson, P., Molloy, E. & Mahoney, P. (2019). Conditions that enable effective feedback. *Higher Education Research & Development,* 38(7), 1401-1416. https://doi.org/10.1080/07294360.2019.1657807

Black, P. & Wiliam, D. (1998). Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan*, 92(1), 81-90. https://doi.org/10.1177/003172171009200119

Hinton, T. (2014). *Impact Management Planning and Evaluation Ladder*. Potts Point, Australia: Abound Consulting.

Lave, J., & Wenger, E. (1990). *Situated Learning: Legitimate Peripheral Participation*. Cambridge, UK: Cambridge University Press.

Malecka, B., Boud, D. & Carless, D. (2020). Eliciting, processing and enacting feedback: mechanisms for embedding student feedback literacy within the curriculum. *Teaching in Higher Education*. https://doi.org/ 10.1080/13562517.2020.1754784

Mercader, C., Ion, G. & Díaz-Vicario, A. (2020). Factors influencing students' peer feedback uptake: instructional design matters. *Assessment & Evaluation in Higher Education*, https://doi.org/10.1080/02602938.2020.1726283

Popham, J. (2005). *Classroom Assessment: What Teachers Need to Know*. Boston: Pearson.

Ramsden, P. (2003). *Learning to Teach in Higher Education*. London & New York: Routledge Falmer. https://doi.org/10.4324/9780203413937

Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, 34(2), 159–179. https://doi.org/10.1080/02602930801956059

Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, 30(2), 175 - 194. https://doi.org/10.1080/0260293042000264262

Sadler, D. R. (1987). Specifying and promulgating achievement standards. *Oxford Review of Education,* 13, 191-209. https://doi.org/10.1080/0305498870130207

Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, 35(5), 535-550. https://doi.org/10.1080/02602930903541015

Sadler. D. R. (2011). Academic freedom, achievement standards and professional identity. *Quality in Higher Education*, 17(1), 85-100, https://doi.org/10.1080/13538322.2011.554639

Wenger, E. (1998). Communities of practice: Learning as a social system. *Systems Thinker,* 9(5), 2-3.

Winstone, N., Bourne, J., Medland, E., Niculescu, I. & Rees, R. (2020). "Check the grade, log out": students' engagement with feedback in learning management systems, *Assessment & Evaluation in Higher Education*. https://doi.org/10.1080/02602938.2020.1787331

Winstone, N. & Boud, D. (2020). The need to disentangle assessment and feedback in higher education. *Studies in Higher Education*. https://doi.org/10.1080/03075079.2020.1779687