

Consistency, Standards and Complexity in Academic Assessment Integrity: Findings from Staff and Student Focus Groups

Miranda Daly, Paul J. Glew and Diana Jefferies

Western Sydney University, Australia

Abstract

This research explores the critical intersection of consistency, standards and complexity in academic assessment integrity using focus groups consisting of academic staff and undergraduate nursing students. While academic quality standards are mandated by accreditation bodies to ensure accountability and skill acquisition, the marking process remains a complex, intuitive task heavily reliant on a marker's experience and practicality. Significant challenges arise when marking becomes overly criteria-driven at the expense of core assessment goals, leading to variability in judging relevant student ideas or specific technical skills such as referencing. Such variability undermines the reliability and objectivity essential for assessment fairness. To mitigate these risks, pre-marking meetings are identified as essential for establishing shared requirements and expectations among staff. Strategic timing of these supports, ideally pre-semester, provides the necessary framework for quality assurance and enables markers to make meaningful and accurate judgements regarding a student's true knowledge and performance.

Keywords: Focus group; marking; educational research; nursing research; assessment.

Introduction

Marking or moderation of assessments is a critical aspect of tertiary education (Adie et al., 2013; Hasan & Jones, 2024). Marks achieved reflect learning and course progression and determine outcomes for graduation (Curtin University, 2021; Smith & Coombe, 2006). Examinations and marked assessments can comprise as much as 90% of the coursework in an undergraduate degree (Lombardi, 2008). It is therefore vital that marking, as the primary moderator for any assessment, is conducted well and fairly, with no bias or subjectivity by those marking (Erturk et al., 2022).

Literature indicates that for most academics, marking is viewed as a mechanical process that often takes considerable time (Goos et al., 2011; Schwab et al., 2024). Hence, regardless of marking importance, the task of marking is often not one that is appreciated by many academics (Schwab et al., 2024). Despite the use of objective marking rubrics, many markers express concerns regarding subjectivity within marking (Bloxham et al., 2016; Koriati et al., 2020; Orr, 2010; Sætre, 2025).



Except where otherwise noted, content in this journal is licensed under a [Creative Commons Attribution 4.0 International Licence](https://creativecommons.org/licenses/by/4.0/). As an open access journal, articles are free to use with proper attribution. ISSN: 2205-0795

Background

Assessment determines what is deemed important for students as it is a measurement of understanding and development (Maclellan, 2001; Victoria State Government, n.d.), while research also finds that naturally, students focus on what is required to pass the subject (Agustina, 2025; Brown, 2001). Standards need to be met and maintained to ensure quality assurance (Al-Amri et al., 2020; Meader et al., 2014), and objectivity and reliability need to be established to ensure that both the assessment task and the marking is valid (Sa et al., 2019; Zahra et al., 2017). This requires a standardised approach (Singh, 2012; Tertiary Education Quality and Standards Agency [TEQSA], n.d.) wherein marking criteria are applied as a constant to ensure objectivity and therefore reliability (Baker et al., 2021; Street & Gilmore, 1983). Critical to quality control are training, support and ongoing development for markers (Brookhart, 1993; Tertiary Education Quality and Standards Agency [TEQSA], n.d.), with marker training viewed as one of the most important quality assurance requirements (Cashin et al., 2017).

Aim and Context of the Research

The research discussed in this article was part of a larger research study exploring the experiences of nursing academic staff and undergraduate nursing students regarding assessment marking. The larger study consisted of three phases, with the subject of this paper being Phase 3, focus group discussions.

Phase 1 of this study consisted of a staff survey. The survey questions were based on a review of the literature regarding marking. Staff who completed the survey were then invited to participate in Phase 2, in which student participants were also invited. Phase 2 consisted of interviews with pairs of staff and separately with pairs of students. Phase 3 of this research study, and the subject of this paper, consisted of focus group discussions in which staff and students were combined.

The aim of these focus group discussions was to further explore examples of best practice from staff and student perspectives regarding assessment marking in order to advocate for the consistent use of such practices.

Ethics approval to conduct the three phases of the research project was obtained from the Western Sydney University Human Research Ethics Committee, H13262.

Methods

This study was conducted in a large tertiary school of nursing and midwifery across four metropolitan campuses in Sydney, Australia. The study sought to ensure a balance of perceptions by consulting all those involved with the marking process, gathering insights into their experiences to discuss what worked and what needed improvement. Consulting staff and students ensured the aspects of best practice, transparency and fairness were addressed. Staff who coordinated or had coordinated subjects were invited to participate in focus group discussions as they were staff who had created assessments, advised on expectations, and provided peer support and guidance for marking. They had also provided advice for students on assessments, managed enquiries and moderated marking.

Student perspectives were also sought to contribute insights in focus group discussions into what they experienced and appreciated in the marking of their assessments. Student participants ($n=3$) were all undergraduate nursing students; one second year student and two third year students. Staff participants ($n=7$) were permanent or sessional (contracted) nursing academics; two had retired and continued as sessional or contracted staff. See Table 1 for participant details.

Table 1*Participant Details Using Pseudonyms*

Pseudonym	Staff/Student	Background	Year Academic Experience/Study
Rahnee	Staff	Previous sessional, SC, marker, tutor	10
William	Staff	Previous sessional, SC, marker, tutor	14
Capri	Staff	Previous sessional, SC, marker, tutor	13
Natasha	Staff	Previous sessional, SC, marker, tutor	25
Alicia	Staff	Previous sessional, SC, marker, tutor	20
Asha	Staff	Previous permanent, SC, marker, now sessional	35
Denise	Staff	Previous permanent, SC, marker, now sessional	40
Cali	Student	Mature-aged student, 30 years in IT, HR, Bachelor of Teaching	Year 2
Travis	Student	Mature-aged student, Bachelor of Science, Bachelor of Business & Commerce, Bachelor of Applied Leadership & Critical Thinking.	Year 3
Thea	Student	Mature-aged student	Year 3

Note. SC = subject coordinator; IT = information technology; HR = human resources

Student participants were invited through their previous participation in Phase 2, as outlined above. All provided consent and were advised that participation was voluntary, that they would be de-identified and that they could withdraw at any time. Confidentiality was discussed, given the group could not hide participants from each other, as well as the sensitivity of the material and that this was not to be discussed outside of the focus group.

Focus groups were selected with marking as the area of shared experience, a commonality that provided important discussion components (Doria et al., 2018). The focus group process supported the group dynamics, allowing for more spontaneous discussions (Tausch & Menold, 2016).

Historically, focus groups have been used for market research and in healthcare settings (Tausch & Menold, 2016). Focus groups are semi-structured with questions developed by the researcher which focus on the area of interest (Tausch & Menold, 2016). Whilst more difficult to organise due to having a larger group of participants, focus groups can generate more discussion than individual interviews (Tausch & Menold, 2016). The purpose of these focus groups was to facilitate discussion involving permanent and sessional staff and students' perspectives and experiences on marking assessments, focussing on what worked well and how this could be implemented in the future.

The thematic structure was informed by a preliminary literature review which underpinned the focus of the study (Braun & Clarke, 2021). The research team analysed the findings through discussion and collective review. This approach ensured that the data analysis aligned with the academic context of the investigation. The themes that were identified involved Quality and Standards, Consistency and Variability, Practices and Beliefs, and Training and Support.

Questions

The questions addressed the experiences of using examples of assessment responses either prior to or following the submission and the use of marking rubrics, which balanced institutional consistency with the subjective nature of professional judgment. Additionally, the analysis examined how experiences with feedback and technical constraints, such as word counts, impacted the perceived fairness and reliability of assessment outcomes. Finally, the role of ongoing supports and marking meetings was discussed as a critical mechanism for aligning staff expectations and ensuring the pedagogical integrity of the marking process.

Thematic Analysis and Reliability

A deductive approach (Azungah, 2018) was used to analyse the preliminary literature review and to identify the four themes previously mentioned. Each of these themes is outlined in separate headings below, and the discussion of each is illustrated by direct participant quotes from the focus groups.

The research team analysed findings through discussion and collective review ensuring the data analysis remained aligned with the academic context of the investigation.

Balanced perspectives were derived through consulting academic staff and undergraduate nursing students to ensure transparency, fairness and a comprehensive understanding of the marking process. Methodological rigour (Jiao, 2026) helped generate in-depth data through the group discussion and a safe space environment was maintained to ensure participants could provide honest and insightful responses. The following themes contain deidentified student and staff participant quotes obtained during the focus groups.

Theme 1 — Quality and Standards

Quality and standards are agreed upon by accreditation bodies and support accountability, skill acquisition requirements, and quality teaching and learning (Al-Amri et al., 2020). Marking is complex and relies upon experience, intuition and practicality (Brooks, 2012; Hasan & Jones, 2024). However, some consider that:

Marking ... is incredibly undervalued, it's as if it's anyone can do it and it's an easy thing to do, it's not easy, it takes probably nearly twice as long as the time that you've allocated, or ... are paid for. (Asha — staff)

Participants discussed the legitimacy of adhering strictly to the marking criteria, or using the marking rubric, reflecting on the requirement for markers to use their own judgement:

You have to move away from the criteria and ask yourself whether the students answered the question ... I think a really able student can put an essay together in an unusually expected way ... I'd say to ... markers, please use your judgement, because a really accomplished writer will answer this question in a way that the criteria doesn't, because the criteria is linear, but a really good writer can integrate and answer the question. (Denise — staff)

Literature is also supportive of looking beyond the marking rubric, given that students may detail relevant ideas and information that are not in the marking criteria, reiterating the importance of a marker to be able to judge this accordingly (Hasan & Jones, 2024; Rowntree, 1996). This requires the marker to have an adequate level of experience and confidence, which must be supported by the subject coordinator (SC) with training and supports integral to this decision-making process. Crisp (2013) and Smith and Coombe (2006) also found that, too often, student learning, and those who are marking students' work, are driven by the marking criteria rather than by the assessment itself:

Marking criteria, it almost sets you up, you have to write to the criteria, not actually respond to the question that's being asked. (Thea — student)

QuickMarks, which are predetermined marking shortcuts created by the SC, can be used to expedite marking, which leaves more time to provide one-to-one feedback (Turnitin, n.d.). Some staff stated, however, that the "QuickMarks cookie cutter approach does not necessarily work" (Capri — staff). It appeared to be less personal and that at times:

People need to understand ... the purpose of QuickMarks and how to use them effectively ... not just cover a student's assessment with them. (Capri — staff)

However, students noted that they had spent time on their work and appreciated it when a marker had clearly spent adequate individualised time marking their work (Thea & Cali — students).

Notably, referencing was raised as a significant variant, with research also indicating this (Calver, 2015; Mbutho & Hutchings, 2019; Santini, 2018; Singh et al., 2020). Rahnee (staff) remarked that some markers clearly had no idea themselves about referencing. Alicia (staff) also raised this as a concern:

Referencing was pretty bad ... there were markers that didn't do it at all and would allocate high marks. When one student is getting full marks, and one student has referencing errors [and lose marks] ... they lose trust in us.

Some universities do not allow staff to mark their own students. To ensure fairness, random allocation is viewed as best practice so the assessment is marked without prejudice and is judged solely on the marking guide and what has been presented (Crisp, 2013; Sadler, 2005).

One staff member explained:

So, there's also different motivators as well as to why people fail, or why people pass ... students, like when you've got your 50th assessment, for example. You're going to pass them if you mark your own. (Rahnee — staff)

Theme 2 — Consistency and Variability

Consistency ensures due process and effective pedagogy (Nixon et al., 2017); variability impacts reliability and objectivity. Reliability and objectivity are critical in assuring validity and maintenance of standards for marking and assessment outcomes and contribute to assessment fairness and integrity (Sa et al., 2019). In combination, this supports quality assurance measures, providing solid foundations to ensure meaningful and accurate judgements about a student's knowledge, skills, and performance (Crisp, 2010; Henderson et al., 2022).

For staff, early pre-marking meetings were deemed essential to establish consistency, requirements and expectations (Denise — staff). Timing of these meetings was considered critical, ideally occurring pre-semester. In cases where students had questions, staff stated:

I often felt scared what you could say to answer questions from students ... when they had questions about the assessment if I didn't know what was going to be required ... sometimes you would give them some advice and then it would change at the time of the marking. (William — staff)

Notably, regardless of planning and attempted rigour and consistency, markers would do as they chose when marking (Capri & Rahnee — staff). This occurred particularly when high volumes of marking were required. Marking can be intuitive and subjective (Beutel et al., 2017), with markers bringing varying perspectives and experiences that may influence what they look for within the assessment (Bloxham, 2012; Koriati et al., 2020; Orr, 2010). Knowing this, Capri (staff), as an SC, expected all enquiries to be redirected to her to ensure a standard response and to reduce variability:

I know ... tutors run their own show and provide information ... it's hard to manage a small business ... that's what 40 staff are, it's a small business. We have to ensure fairness in processes. (Capri - staff).

Some SCs ensured the requirements were well detailed for the teaching staff pre-semester (Alicia — staff). In addition, to educate her markers, Alicia (staff) would:

Send out the marking guide. Once we've gone over the marking guide and we've agreed on it, I send out a paper ... "mark it completely; I don't want to see just what mark you gave it, I want you to annotate this paper, correct it". ... I know that sounds horrible and it's like micro-managing my staff, but I need to know what the marking is.

Variability by those marking was obvious to some students, with some markers clearly taking more time and care. For example, both Travis and Thea (students) talked of "good and bad markers" with some putting more effort than others into marking students' work.

Participants observed that assessment task word counts must be sufficient to include the required points:

When you write an assessment item ... you should write the expected response. I think it needs to be reviewed by ... other people. You know, to write the response that you expect to the question, so ... does that fit in with the prescribed word limit and ... does that expected response really address the question ... because I think you need to be very, very clear. That review is essential. (Natasha — staff)

Alicia (staff) detailed potential responses to her markers, indicating that she knew the expected answers were achievable within the word constraints:

I always have scripted answers, so I know that I'm not asking students for an impossible request. I also indicate ... approximately how many words in this particular short answer question. ... I do the assessment with the deputy subject coordinator and ... more experienced staff. I also include in my marking some sort of framework ... "here are some other comments that the students might provide".

The SC guaranteeing that the assessment is feasible within the word constraints ensures achievability, as many students do not have the necessary skills to articulate their knowledge in a concise academic manner to communicate a succinct and correct response.

Capri (staff) noted that word limits preclude most students from providing complete answers:

Is there enough to answer this assessment, and the answer is actually no, it's not and I know it's not, unless you're extremely articulate which is you know ... 2% of the students who we are catering for ... we're asking to write about the pathophysiology It's the constraints ... so asking them about the pathophysiology and how it relates to that patient in the case study But to answer that in 1250 words means you miss stuff, unless you're extremely articulate right?

She also commented that if students were unsure of how to articulate their explanation within the required assessment word limit, they were more likely to resort to looking for alternatives:

If the student can't fit it in the words they'll ... copy and paste, they'll take it from other students, they will do anything ... if they're lacking the model to work with, they're at a complete loss, and so they end up going down the path of academic misconduct, because it's a last resort of desperation. (Capri — staff)

For students, there was much agonising at times as they were unsure what to add or leave out when there were word count restrictions:

You ... have to think, how much am I going to sacrifice for each section, so if that's worth 30, that's worth 10 okay, I might not have a very good thesis statement in my introduction and repeat it in my conclusion because I need more words in the body paragraphs. If it goes over the word count the 1500 and you've got 1650 or whatever it is ... the marker will stop. And if you're over that you're penalised, and I have experienced that. (Thea — student).

Travis, another student, expressed his frustration at being asked in the feedback to provide more details when this was just not possible:

It's all destroying when you don't get part of your assignment read because you're over the word count, but when you do cut things and then you get QuickMarks that says, "you should have said this", I am like, "I cut that out due to the word count!"

Theme 3 — Practices and Beliefs

Marking practices stem from a marker's knowledge, values, integrity, behaviour, attention to detail, and respect (Singh et al., 2020). Beliefs develop through practical experience, particularly when staff actively teach and assess in a discipline, as diverse pedagogical perspectives shape content mastery and assessment understanding (Herppich et al., 2010; Latif & Wasim, 2022). In turn, these experience-grounded beliefs directly drive marking practices by framing how markers interpret standards and evaluate student work, while repeated marking practice continuously refines and perfects their skill (Singh et al., 2020).

For some staff and students, being shown what was required helped them towards the goal of answering the question. However, being shown after submission resulted in some students feeling "like I was stupid because I did not reach that level" (Thea — student) or "Oh my God there's no way I could achieve that ... oh my gosh I've left this out this, this, and this" (Cali — student).

Travis (student) explained, "I have no idea what I'm doing half the time and I'm using those exemplars [if provided early]". However, Capri (staff) noted that exemplars were challenging to write if a previous example was not available:

It's great if you can rely on a previous student who has done great work that you can show again, but usually, we have to change the assessments so that is not always possible So, for someone to sit and write is ... time-consuming.

Handley et al. (2013) indicated that the more complex an assessment task, the more examples should ideally be provided as a guide for both staff and students to provide a reference point to inform judgement (NSW Government, 2025; To et al., 2022). Asha (staff) noted:

Because I would have thought that the plan is how to give students information about how they should approach the topic, rather than afterwards, going, "this is what you should have done".

Theme 4 — Training and Support

Feedback is seen as a further teachable moment (Hadjieconomou & Tombs, 2021; O'Donovan et al., 2021; Walvoord & Anderson, 2011). Participants discussed the pedagogical potential of feedback:

When providing written feedback ... it was an opportunity to personalise it. (Capri — staff)

From the student's perspective, Travis stated, "I don't look at a mark, I look at the feedback first". Thea (student) explained that adequate feedback lets you "learn, develop and grow". It allowed her to make sense of what she had done, or not done:

Feedback is an opportunity to continue your learning, not so much what I've done wrong ... but what I've done right, what's good, what's not so good ... why and how, and where we can go. (Thea — student)

One student participant (Travis) found that he failed a particular assignment despite using the same style of writing he had used for other assessments (over a year), including his references. Through the feedback, he realised that what he had previously presented had, in fact, been incorrect:

I went from a high distinction to a credit, but every time a mark was taken off it was justified, it was explained, that was feedback ... with comments I knew exactly what I had to do better ... when it was the other subject, I had to figure it out, as no feedback.

There is a presumption that all students understand what is required of them as well as academic language, subject aims, outcomes and criteria (Deeley et al., 2019; Hadjieconomou & Tombs, 2021). Students have varying degrees of comprehension and ability; therefore, varying marks are awarded:

Because it's not always achievable for everyone to be at the top, simple as that. (Capri — staff)

Natasha (staff) felt that, at times, it was essential to look further and provide feedback on all the information provided, areas that were good and areas that needed improvement:

It is really important to not just focus on what is not there, but also to reward or point out good content or good aspects of a response, I think it can be quite soul destroying if a student looks at the paper ... a lot of the feedback is negative, and I think sometimes we need to look for good points and point that out.

Positive feedback encourages students and makes them more likely to absorb the responses, thereby increasing their motivation (Henderson et al., 2019; O'Donovan et al., 2021; Walvoord & Anderson, 2011; Wimshurst & Manning, 2013).

Asha (staff) felt strongly that feedback needed to be done well with feedback needing to be:

So relevant to them [the students] that it's a feed forward ... to know why these things are important in this context. When you work with them ... you're having an interaction ... that connection.

Denise (staff) also noted that "Respectfully given feedback I think they appreciate". Feedback is used to "join the dots" for the student (Asha — staff).

Thea (student) fully appreciated feedback that appeared to address her and her work:

I love comments ... I feel like I'm not just another number ... I'm a person who's ... spent a lot of time doing the assignment and I know that the marker spent a lot of time marking ... and that just tells me that they actually engaged in marking this ... and it's not just another tick box exercise.

Asha (staff) found that, at times, students begged for a review or more marks, so detailed feedback was important to contextualise the marks awarded in a way that the student could understand:

There's more justification required to fail ... so you really have to give good feedback.

Staff mentioned that some markers focused on grammar and punctuation rather than the content of the assessment. Some staff saw this as marking the "easy pickings" and "lazy marking" (Denise & Asha). This also highlights the importance of familiarity with the content and context of the unit. For this reason, several staff felt it was important that those teaching in the subject also marked (Natasha, Alicia, Denise, Asha — staff).

The whole purpose of feedback is that the student reads it and learns from it. Two staff members strongly agreed, even to the extent of using the same words, that unfortunately this is not always the case:

If you're not reading the feedback, it defeats the purpose of even having it provided in the first place. (Denise & Asha)

Notably, some students tend to focus on the mark, rather than the educational value of written comments (Bailey & Garner, 2010; O'Keeffe et al., 2018; Pitt & Norton, 2017).

The benefits of training and support for markers prior to commencing this important task is clear (Chen et al., 2017; Norton et al., 2019), with the use of ongoing supports being positively linked to job satisfaction, retention and stability for sessional markers (Cowin & Moroney, 2018). This can also help with quality assurance and meeting the standards of a regulatory framework (Al-Amri et al., 2020; Crimmins et al., 2016).

Some SCs would request a few papers to be marked initially for them to review and provide one-to-one feedback to the marker. This improved quality assurance and consistency and helped the marker to work in accordance with expectations. Alicia (staff) commented that:

Quality control is an important aspect of marking, but it is also time consuming.

Time, workloads and costs are major considerations and impact on assessment practices for many staff and facilities (Fry et al., 2015; Richards & Richardson, 2019; Smith & Coombe, 2006; Walvoord & Anderson, 2011). Table 2 summarises the core focus and key concepts relating to each of the four themes.

Table 2

Themes, Core Focus and Key Concepts

Theme	Core Focus and Key Concepts
Quality and Standards	Institutional accountability, accreditation and the use of professional judgment to look "beyond the criteria" of linear rubrics
Consistency and Variability	The essential role of pre-marking meetings, the impact of the "referencing gap" and how technical constraints like word counts affect fairness
Practices and Beliefs	The influence of a marker's values and integrity on conduct, including the strategic use of exemplars to guide students
Training and Support	Utilising feedback as a "teachable moment" (feed forward) and the importance of ongoing support for marker consistency and job satisfaction

Discussion

The marking process is fundamentally an intuitive and subjective task that relies on a marker's experience and practicality (Beutel et al., 2017; Bloxham, 2009). However, a significant challenge arises when marking becomes overly driven by rigid criteria at the expense of the assessment's core objectives (Koriat et al., 2020). To ensure quality control, markers must often exercise professional judgement to evaluate relevant student ideas or information that may not be explicitly captured within a rubric (Bloxham, 2012; Chan & Ho, 2019). This can be challenging, with markers bringing varying degrees of experience.

Ensuring consistency is critical for the validity and maintenance of standards, which are ultimately overseen by accreditation bodies to ensure institutional accountability and quality teaching (Al-Amri et al., 2020; Australian Nursing & Midwifery Accreditation Council [ANMAC], 2018).

The main findings from this study can be summarised as follows:

1. The Paradox of "Beyond the Rubric" Judgment

There exists critical tension between rigid criteria and professional intuition (Koriat et al., 2020). While rubrics aim for objectivity, experienced markers found it necessary to move "beyond the criteria" to accurately judge if a student had truly

answered the question, particularly when a high-achieving student integrated ideas in a less conventional way. This highlights that marking is a complex, intuitive task rather than a purely mechanical one (Crimmins et al., 2016; Sa et al., 2019) which requires a level of confidence and experience and quality control oversight that needs to be supported by SCs.

2. Technical Constraints as a Catalyst for Academic Misconduct

A high-impact finding is the link between restrictive word counts and academic integrity risks.

- **Desperation and Misconduct:** When word limits were too tight for complex subjects like pathophysiology, students stated that they were forced to make choices and "sacrifice" essential sections or, as staff reported, resort to academic misconduct (e.g., copying and pasting) as a "last resort of desperation".
- **The Clarity Gap:** Staff emphasised that word limits must be tested against an "expected response" to ensure the task is achievable for the average student, not just the top 2% of students who are "extremely articulate".

3. The "Referencing Gap" and Loss of Student Trust

Referencing was identified as a major source of unacceptable variability in marking (Singh et al., 2020).

- **Marker Inconsistency:** Some staff members lacked a clear understanding of referencing themselves (Hasan & Jones, 2024), leading to marker variations whereby some students lost marks from one marker, yet others obtained full marks for the same errors from another.
- **Impact on Credibility:** This discrepancy caused students to "lose trust" in the academic system and underscores the need for standardised referencing training for staff.

4. Strategic Timing of Quality Assurance (Pre-Semester vs. Post-Submission)

The timing of support mechanisms was crucial for maintaining standards:

- **Proactive vs. Reactive:** Pre-marking meetings are essential but must occur pre-semester to ensure staff give students consistent advice *before* they submit (Sætre, 2025).
- **Exemplar Effectiveness:** While students valued exemplars for clarity, providing them only *after* submission could be "soul destroying", making students feel "stupid" for what they missed, rather than aiding their learning.

5. Feedback as "Feed Forward" and Personal Connection

Feedback is most impactful when it functions as a "teachable moment" rather than just a justification for a grade (Hadjieconomou & Tombs, 2021; O'Donovan et al., 2021; Walvoord & Anderson, 2011).

- **Individualisation:** Students highly valued personalised feedback and felt like "just another number" when markers relied solely on automated "cookie-cutter" tools like QuickMarks.
- **Motivational Impact:** Constructive feedback pointing out "what's right" increased student motivation and helped them "join the dots" of their learning, provided they had actually read the feedback.

Systemic Recommendations for Best Practice

This study advocates for specific systemic changes to bridge the gap between intuition and rigid criteria (Crisp, 2013; Sadler, 2005):

- **Random Allocation:** To ensure fairness and remove personal bias or "motivators" (such as passing a student just to finish their allocated workload), randomly allocating assessments, rather than having tutors mark their own students, is recommended as best practice.
- **Undervalued Labor:** There was a strong consensus among staff that marking is incredibly undervalued and that it often took twice as long as the time allocated or paid, which could lead to "lazy marking" and a focus on "easy pickings" such as grammar instead of substantive content.

Strengths and Limitations of This Study

As larger groups allow a variety of conversations to occur, focus groups were utilised to help inform decisions in a consultative process (Doria et al., 2018). This was clear during the discussions wherein staff and students were all able to have a voice and discuss their experience and ideals for marking.

Having more people in an interview reduces confidentiality and anonymity, and the dynamics may be less predictable with the researcher having less control; hence, informed consent with full disclosure for participants is required to ensure all are aware of the process (Mwita, 2022; Sim & Waterfield, 2019). Careful consideration was required by the researchers to ensure the environment remained a safe space, with procedures in place to ensure all participants felt comfortable. The conversations flowed well and were productive and insightful for all who attended. Whilst conducted across multiple campuses, these focus groups were conducted in one school only; there may have been opportunities to capture variations in responses if the groups spanned across other specialties or universities.

Conclusion

As this study was conducted within a single school of nursing, its findings may be influenced by the specific clinical and technical requirements of that field. Future research should explore whether the tension between professional intuition and rigid rubrics is consistent across other disciplines, such as the humanities or creative arts, where integrated responses might be even more common.

While both staff and students identified feedback as a vital "teachable moment" for growth, students often focus on the mark rather than the comments. Future research could longitudinally track student performance to determine if specific "feed-forward" strategies (e.g., personalised interactions vs. automated tools like QuickMarks) result in measurable improvements in subsequent assessment tasks.

A significant finding in this study was that restrictive word counts for complex subjects can drive students toward academic misconduct through desperation. Future research should investigate the correlation between assessment design (specifically word limits) and integrity breaches to establish evidence-based guidelines for "achievable" word counts in high-complexity subjects like pathophysiology.

Given that referencing was identified as a major source of variability—often stemming from a lack of staff expertise—future research could evaluate the effectiveness of mandatory, standardised marker training programs. Research could also explore whether automating technical checks (like referencing) allows markers to spend more time on substantive content, thereby reducing marking focused on areas such as grammar only.

Future research should examine the relationship between marker compensation/workload models and the quality of feedback. Understanding how systemic pressures influence a marker's decision to engage deeply with content versus performing a "tick box exercise" would be crucial for institutional policy development.

The gap between rigid criteria and professional intuition must be bridged. Future research could focus on developing training frameworks that teach markers how to apply professional judgment consistently. This could involve studying the effectiveness of "consensus marking" or pre-marking "annotation exercises" in reducing subjectivity and building marker confidence.

Conflict of Interest

No conflict of interest has been declared by the authors.

AI Use

During the preparation of this work the author(s) used LM Notebook (Google, 2024). to assist with reducing the word count to fit the submission into the required limit. This is the authors' original work, it is part of the authors PhD research findings, with timelines available for reviews undertaken and progression evidence. Prior to and after using this tool/service, the author(s) reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

- Adie, L., Lloyd, M., & Beutel, D. (2013). Identifying discourses of moderation in higher education. *Assessment & Evaluation in Higher Education*, 38(8), 968–977. <https://doi.org/10.1080/02602938.2013.769200>
- Agustina, D. (2025). *Assessment strategies used by lecturer in teaching critical reading* [Undergraduate thesis, Institut Agama Islam Negeri Curup]. E-Theses IAIN Curup. <http://e-theses.iaincurup.ac.id/id/eprint/9236>
- Al-Amri, A. S., Mathew, P., Zubairi, Y. Z., & Jani, R. (2020). Optimal standards to measure the quality of higher education institutions in Oman: Stakeholders' perception. *SAGE Open*, 10(3). <https://doi.org/10.1177/2158244020947440>
- Australian Nursing & Midwifery Accreditation Council. (n.d.). *Accreditation Standards*. ANMAC <https://www.anmac.org.au/accreditation-standards>
- Azungah, T. (2018). Qualitative research: Deductive and inductive approaches to data analysis. *Qualitative Research Journal*, 18(4), 383–400. <https://doi.org/10.1108/QRJ-D-18-00035>
- Bailey, R., & Garner, M. (2010). Is the feedback in higher education assessment worth the paper it is written on? Teachers' reflections on their practices. *Teaching in Higher Education*, 15(2), 187–198. <https://doi.org/10.1080/13562511003620019>
- Baker, C., Cary, A. H., & da Conceicao Bento, M. (2021). Global standards for professional nursing education: The time is now. *Journal of Professional Nursing*, 37(1), 86–92. <https://doi.org/10.1016/j.profnurs.2020.10.001>
- Beutel, D., Adie, L., & Lloyd, M. (2017). Assessment moderation in an Australian context: Processes, practices, and challenges. *Teaching in Higher Education*, 22(1), 1–14. <https://doi.org/10.1080/13562517.2016.1213232>
- Bloxham, S. (2009). Marking and moderation in the UK: False assumptions and wasted resources. *Assessment & Evaluation in Higher Education*, 34(2), 209–220. <https://doi.org/10.1080/02602930801955978>
- Bloxham, S. (2012). 'You can see the quality in front of your eyes': Grounding academic standards between rationality and interpretation. *Quality in Higher Education*, 18(2), 185–204. <https://doi.org/10.1080/13538322.2012.711071>
- Bloxham, S., den-Outer, B., Hudson, J., & Price, M. (2016). Let's stop the pretence of consistent marking: Exploring the multiple limitations of assessment criteria. *Assessment & Evaluation in Higher Education*, 41(3), 466–481. <https://doi.org/10.1080/02602938.2015.1024607>
- Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, 18(3), 328–352. <https://doi.org/10.1080/14780887.2020.1769238>
- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement*, 30(2), 123–142. <https://doi.org/10.1111/j.1745-3984.1993.tb01070.x>
- Brooks, V. (2012). Marking as judgment. *Research Papers in Education*, 27(1), 63–80. <https://doi.org/10.1080/02671520903331008>
- Brown, G. (2001). *Assessment: A guide for lecturers*. LTSN Generic Centre. https://www.researchgate.net/publication/265398364_Assessment_A_Guide_for_Lecturers
- Calver, M. (2015). The importance of authors ensuring referencing and page proofs are correct. *Pacific Conservation Biology*, 21(3), 173–174. https://doi.org/10.1071/PCv21n3_ED
- Cashin, A., Heartfield, M., Bryce, J., Devey, L., Buckley, T., Cox, D., Kerdo, E., Kelly, J., Thoms, D., & Fisher, M. (2017). Standards for practice for registered nurses in Australia. *Collegian*, 24(3), 255–266. <https://doi.org/10.1016/j.colegn.2016.03.002>
- Chan, Z., & Ho, S. (2019). Good and bad practices in rubrics: The perspectives of students and educators. *Assessment & Evaluation in Higher Education*, 44(4), 533–545. <https://doi.org/10.1080/02602938.2018.1522528>
- Chen, C.-Y., Chang, H., Hsu, W.-C., & Sheen, G.-J. (2017). Learning, behaviour and reaction framework: a model for training raters to improve assessment quality. *Assessment & Evaluation in Higher Education*, 42(5), 705–723. <https://doi.org/10.1080/02602938.2016.1180663>
- Cowin, L. S., & Moroney, R. (2018). Modelling job support, job fit, job role and job satisfaction for school of nursing sessional academic staff. *BMC nursing*, 17(1), Article 22. <https://doi.org/10.1186/s12912-018-0290-2>
- Crimmins, G., Nash, G., Oprescu, F., Alla, K., Brock, G., Hickson-Jamieson, B., & Noakes, C. (2016). Can a systematic assessment moderation process assure the quality and integrity of assessment practice while supporting the professional development of casual academics? *Assessment & Evaluation in Higher Education*, 41(3), 427–441. <https://doi.org/10.1080/02602938.2015.1017754>
- Crisp, V. (2010). Towards a model of the judgement processes involved in examination marking. *Oxford Review of Education*, 36(1), 1–21. <https://doi.org/10.1080/03054980903454181>
- Crisp, V. (2013). Criteria, comparison and past experiences: How do teachers make judgements when marking coursework? *Assessment in Education: Principles, Policy & Practice*, 20(1), 127–144. <https://doi.org/10.1080/0969594X.2012.741059>
- Curtin University. (2018). *Assessment and Student Progression*. https://s43414.pcdn.co/student/wp-content/uploads/sites/3/2024/09/Assessment_and_Student_Progression_Manual.pdf

- Deeley, S. J., Fischbacher-Smith, M., Karadzhev, D., & Koristashevskaya, E. (2019). Exploring the ‘wicked’ problem of student dissatisfaction with assessment and feedback in higher education. *Higher Education Pedagogies*, 4(1), 385–405. <https://doi.org/10.1080/23752696.2019.1644659>
- Doria, N., Condran, B., Boulos, L., Maillet, D. G. C., Dowling, L., & Levy, A. (2018). Sharpening the focus: Differentiating between focus groups for patient engagement vs. qualitative research. *Research Involvement and Engagement*, 4(1), Article 19. <https://doi.org/10.1186/s40900-018-0102-6>
- Erturk, S., van Tilburg, W. A. P., & Igou, E. R. (2022). Off the mark: Repetitive marking undermines essay evaluations due to boredom. *Motivation and Emotion*, 46(2), 264–275. <https://doi.org/10.1007/s11031-022-09929-2>
- Google. (2024). NotebookLM. In (Vol. AI Tool): NotebookLM.
- Goos, M., Gannaway, D., & Hughes, C. (2011). Assessment as an equity issue in higher education: Comparing the perceptions of first year students, course coordinators, and academic leaders. *The Australian Educational Researcher*, 38(1), 95–107. <https://doi.org/10.1007/s13384-010-0008-2>
- Hadjieconomou, S., & Tombs, M. (2021). Postgraduate students’ perceptions of what makes for effective assessment feedback: A case study of a clinical masters course. *Postgraduate Medical Journal*, 97(1150), 491–494. <https://doi.org/doi:10.1136/postgradmedj-2020-137538>
- Handley, K., den Outer, B., & Price, M. (2013). Learning to mark: Exemplars, dialogue and participation in assessment communities. *Higher Education Research & Development*, 32(6), 888–900. <https://doi.org/10.1080/07294360.2013.806438>
- Hasan, A., & Jones, B. (2024). Assessing the assessors: investigating the process of marking essays. *Frontiers in Oral Health*, 5, Article 1272692. <https://doi.org/10.3389/froh.2024.1272692>
- Henderson, B., Chipchase, L., Aitken, R., & Lewis, L. K. (2022). Consensus marking as a grading method for the development of evaluative judgement: Comparing assessor and students. *Nurse Education in Practice*, 63, Article 103386. <https://doi.org/10.1016/j.nepr.2022.103386>
- Henderson, M., Ryan, T., & Phillips, M. (2019). The challenges of feedback in higher education. *Assessment & Evaluation in Higher Education*, 44(8), 1237–1252. <https://doi.org/10.1080/02602938.2019.1599815>
- Herppich, S., Wittwer, J., Nuckles, M., & Renkl, A. (2010). Do tutors’ content knowledge and beliefs about learning influence their assessment of tutees’ understanding? *Proceedings of the Annual Meeting of the Cognitive Science Society*, 32, 314–319. <https://escholarship.org/uc/item/1c169462>
- Jiao, R. (2026). Demystifying verification and validation in design research methodology - from underdeveloped practice to methodological rigour. *Journal of Engineering Design*, 37(4), 1143–1162. <https://doi.org/10.1080/09544828.2026.2624355>
- Koriat, A., Undorf, M., Newman, E., & Schwarz, N. (2020). Subjective confidence in the response to personality questions: Some insight into the construction of people’s responses to test items. *Frontiers in Psychology*, 11, 1250–1250. <https://doi.org/10.3389/fpsyg.2020.01250>
- Latif, M. W., & Wasim, A. (2022). Teacher beliefs, personal theories and conceptions of assessment literacy—a tertiary EFL perspective. *Language Testing in Asia*, 12(1), Article 11. <https://doi.org/10.1186/s40468-022-00158-5>
- Lombardi, M. M. (2008). *Making the grade: The role of assessment in authentic learning* (D. Oblinger, Ed.). EDUCAUSE Learning Initiative. <https://phongdbcl.ntu.edu.vn/uploads/47/files/old/Tu%20lieu%20tham%20khao/Phuong%20phap%20danh%20gia/authentic%20assessment%203.pdf>
- MacLellan, E. (2001). Assessment for learning: The differing perceptions of tutors and students. *Assessment & Evaluation in Higher Education*, 26(4), 307–318. <https://doi.org/10.1080/02602930120063466>
- Mbutho, N. P., & Hutchings, C. (2019). Factors that influence the ways university students think about referencing and related concepts. *South African Journal of Higher Education*, 33(5), 230–248. <https://doi.org/10.20853/33-5-3575>
- Meador, N., King, K., Llewellyn, A., Norman, G., Brown, J., Rodgers, M., Moe-Byrne, T., Higgins, J. P., Sowden, A., & Stewart, G. (2014). A checklist designed to aid consistency and reproducibility of GRADE assessments: Development and pilot validation. *Systematic Reviews*, 3(1), Article 82. <https://doi.org/10.1186/2046-4053-3-82>
- Mwita, K. (2022). Strengths and weaknesses of qualitative research in social science studies. *International Journal of Research in Business & Social Science*, 11(6), 618–625. <https://dx.doi.org/10.2139/ssrn.4880484>
- Nixon, S., Brooman, S., Murphy, B., & Fearon, D. (2017). Clarity, consistency and communication: Using enhanced dialogue to create a course-based feedback strategy. *Assessment & Evaluation in Higher Education*, 42(5), 812–822. <https://doi.org/10.1080/02602938.2016.1195333>
- Norton, L., Floyd, S., & Norton, B. (2019). Lecturers’ views of assessment design, marking and feedback in higher education: A case for professionalisation? *Assessment and Evaluation in Higher Education*, 44(8), 1209–1221. <https://doi.org/10.1080/02602938.2019.1592110>
- NSW Government. (2026). *Effective assessment advice*. <https://education.nsw.gov.au/teaching-and-learning/assessment/strengthening-assessment/effective-assessment-practice>

- O'Donovan, B. M., den Outer, B., Price, M., & Lloyd, A. (2021). What makes good feedback good? *Studies in Higher Education*, 46(2), 318–329. <https://doi.org/10.1080/03075079.2019.1630812>
- O'Keeffe, M., Gormley, C., & Ferguson, P. B. (2018). Moving the focus from grades to feedback: A case study of pass/fail marking. *Practitioner Research in Higher Education*, 11(1), 70–79. <https://ojs.cumbria.ac.uk/index.php/prhe/article/view/463>
- Orr, S. (2010). We kind of try to merge our own experience with the objectivity of the criteria: The role of connoisseurship and tacit practice in undergraduate fine art assessment. *Art, Design & Communication in Higher Education*, 9(1), 5–19. https://doi.org/10.1386/adch.9.1.5_1
- Pitt, E., & Norton, L. (2017). 'Now that's the feedback I want!' Students' reactions to feedback on graded work and what they do with it. *Assessment & Evaluation in Higher Education*, 42(4), 499–516. <https://psycnet.apa.org/doi/10.1080/02602938.2016.1142500>
- Richards, G., & Richardson, R. (2019). *Reducing teachers' marking workload and developing pupils' learning: How to create more impact with less marking*. Routledge. <https://doi.org/10.4324/9780429243288>
- Rowntree, D. (1987). *Assessing Students: How shall we know them?* (2nd ed.). Kogan Page Pty Ltd.
- Sa, B., Ezenwaka, C., Singh, K., Vuma, S., & Majumder, M. A. A. (2019). Tutor assessment of PBL process: Does tutor variability affect objectivity and reliability? *BMC Medical Education*, 19(1), Article 76. <https://doi.org/10.1186/s12909-019-1508-z>
- 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>
- Sætre, C. (2025). Behind the grades: Co-constructing fairness to reach agreement in evaluative judgement. *Assessment & Evaluation in Higher Education*, 50(2), 236–249. <https://doi.org/10.1080/02602938.2024.2373789>
- Santini, A. (2018). The importance of referencing. *The Journal of Critical Care Medicine*, 4(1), 3–4. <https://doi.org/10.2478/jccm-2018-0002>
- Schwab, C., Frenzel, A. C., Jaeger, J., Lorenz, A. B., & Stupnisky, R. H. (2024). How do university faculty feel about grading? Insights from a control-value theory perspective. *Studies in Higher Education* 49(8), 1486–1503. <https://doi.org/10.1080/03075079.2023.2269190>
- Sim, J., & Waterfield, J. (2019). Focus group methodology: Some ethical challenges. *Quality & Quantity*, 53(6), 3003–3022. <https://doi.org/10.1007/s11135-019-00914-5>
- Singh, D., Steenkamp, M., Harmse, T., & Botha, J. C. (2020). Engaging the student voices to improve referencing skills and practices in higher education: A South African case study. *South African Journal of Higher Education*, 34(5), 122–135. <https://doi.org/10.20853/34-5-4199>
- Singh, T. (2012). Student assessment: Issues and dilemmas regarding objectivity. *The National Medical Journal of India*, 25(5), 287–290. <https://pubmed.ncbi.nlm.nih.gov/23448630/>
- Smith, E., & Coombe, K. (2006). Quality and qualms in the marking of university assignments by sessional staff: An exploratory study. *Higher Education*, 51(1), 45–69. <https://doi.org/10.1007/s10734-004-6376-7>
- Street, M. R., & Gilmore, A. M. (1983). The reliability of 'expert judgement' in the marking of state nursing examinations. *Journal of Advanced Nursing*, 8(3), 221–226. <https://doi.org/10.1111/j.1365-2648.1983.tb00317.x>
- Tausch, A. P., & Menold, N. (2016). Methodological aspects of focus groups in health research: Results of qualitative interviews with focus group moderators. *Global Qualitative Nursing Research*, 3. <https://doi.org/10.1177/2333393616630466>
- Tertiary Education Quality and Standards Agency (TEQSA). (n.d.). *About us*. Retrieved July 8, 2022, from <https://www.teqsa.gov.au/about-us-0>
- Tertiary Education Quality and Standards Agency (TEQSA). (n.d.). *Assessment insights*. <https://www.teqsa.gov.au/guides-resources/resources/sector-updates-and-alerts/assessment-insights>
- To, J., Panadero, E., & Carless, D. (2022). A systematic review of the educational uses and effects of exemplars. *Assessment & Evaluation in Higher Education*, 47(8), 1167–1182. <https://psycnet.apa.org/doi/10.1080/02602938.2021.2011134>
- Turnitin. (n.d.). <https://guides.turnitin.com/hc/en-us/articles/23423780386061--Classic-Using-QuickMarks>
- Victoria State Government. (n.d.). *Professional practice: Assessment moderation*. <https://vcaa.vic.edu.au/assessment/foundation-10/formative-assessment/develop-formative-assessment-rubrics/module-5-moderate>
- Walvoord, B. E., & Anderson, V. J. (2011). *Effective grading: A tool for learning and assessment in college* (2nd ed.). John Wiley & Sons. https://books.google.com.au/books/about/Effective_Grading.html?id=EJxy06yX_NoC&redir_esc=y
- Wimshurst, K., & Manning, M. (2013). Feed-forward assessment, exemplars and peer marking: Evidence of efficacy. *Assessment & Evaluation in Higher Education*, 38(4), 451–465. <https://doi.org/10.1080/02602938.2011.646236>
- Zahra, D., Robinson, I., Roberts, M., Coombes, L., Cockerill, J., & Burr, S. (2017). Rigour in moderation processes is more important than the choice of method. *Assessment & Evaluation in Higher Education*, 42(7), 1159–1167. <https://doi.org/10.1080/02602938.2016.1236183>

Please cite this article as:

Daly, M., Glew, P. J., & Jefferies, D. (2026). Consistency, standards and complexity in academic assessment integrity: Findings from staff and student focus groups. *Student Success*. Advance online publication. <https://doi.org/10.63608/ssj.4076>

This article has been accepted for publication in *Student Success*. Please see the Editorial Policies under the 'About' section of the Journal website for further information.

Student Success: *A journal exploring the experiences of students in tertiary education.*



Except where otherwise noted, content in this journal is licensed under a [Creative Commons Attribution 4.0 International Licence](https://creativecommons.org/licenses/by/4.0/). As an open access journal, articles are free to use with proper attribution. ISSN: 2205-0795