Research article

Understanding the processes, practices and influences of calibration on feedback literacy in higher education marking: A qualitative study ☆

Rebekkah Middleton [a,*], Kelly Lewer [a], Carolyn Antoniou [b], Helen Pratt [a], Suzanne Bowdler [a], Carley Jans [a], Kaye Rolls [a]

[a] School of Nursing, Faculty of Science, Medicine & Health, Australia
[b] School of Nursing and Midwifery, Australia

ARTICLE INFO

ABSTRACT

Background: In the University setting, striving for consistency and reliability of assessment evaluation is essential to reducing the impact of marking variations. Marking processes such as consensus calibration have the potential to reduce issues which arise from the influence of markers professional knowledge and experience, as well as fixed and acquired marking habits. Furthermore, the influence of marker feedback which supports learning development is associated with the feedback literacy of both the teacher and the student. A gap is currently present in the literature as these practices are not discussed together.
Objectives: To explore how nursing academics perceive and understand calibration practices and associated feedback literacy.
Design: Theoretical underpinnings in participatory and person-centred research methodology supported the critical ontological perspective of this study where the intent of the research was to explore the reality that exists within the context where the research was conducted.
Setting: A single School of Nursing in an Australian University with six campuses spanning metropolitan, regional and rural sites.
Participants: Nursing academics and casual tutors with various levels of experience in assessment marking and feedback.
Methods: Semi-structured group interviews that were analysed using reflexive thematic analysis.
Results: Four overarching themes were identified; rubrics, calibration, feedback and justice.
Conclusion: Calibration improves staff cohesion, fosters better practices and consistency, and permits nuanced interpretation of assessments while maintaining uniformity.
Enhanced feedback literacy that integrates principles of equity, justice, and learner-centeredness is required. Fundamentally calibration guides educators toward holistic approaches that foster consistency, equity, and thorough feedback practices.

1. Introduction

Inconsistency and unreliability in marking within university settings largely occurs due to variation in professional knowledge, experience, values and policy (Bloxham et al., 2015; Read et al., 2005). In addition, fixed and acquired marking habits can influence grading in unconscious ways (Bloxham et al., 2015; O'Connell et al., 2016). It is, therefore, important to ensure reliable marking processes occur to address concerns around marking inconsistencies (Boud et al., 2015; Herbert et al., 2014) and ensure alignment to curriculum learning outcomes (O'Connell et al., 2016; Sadler, 2013). In this paper we will present a qualitative study that examines how nursing academics perceived and understood

calibration practices and associated feedback literacy.

To minimise such disparity in higher education academic assessment marking practices and outcomes, it is widely advised that consensus calibration forums occur (Cleary et al., 2019; Bloxham and Price, 2013; O'Connell et al., 2016; Rust, 2009; Sadler, 2013). This process is particularly important when student cohorts are large, as is the case in many pre-registration nursing courses where marking is undertaken by a team of markers (Herbert et al., 2014). Such calibration processes, where markers negotiate a shared meaning that will enable reliable judgement of student responses to occur independently, are critical (Sadler, 2013). They ensure that markers approach the academic assessment in an attuned manner. This is a way to be able to manage variations in student answers, rather than being standardised which acknowledges only an 'ideal' response (Cleary et al., 2019).

Nursing academics, as part of the marking process, provide feedback to students. In the words of Carless and Boud (2018) feedback literacy is the: "understandings, capacities, and dispositions needed to make sense of information and use it to enhance work or learning strategies" (p. 1316). Carless and Winstone (2023, p. 4) built on this definition by incorporating "knowledge, expertise and dispositions to design feedback processes" that facilitate students to understand and apply feedback, thereby advancing their feedback literacy. Quigley (2021) confers that healthcare students feedback literacy needs to be intentionally supported through clear and open measures to increase agency and active engagement with feedback.

When nursing academics approach feedback from a teacher-driven model of 'feedback as telling', they limit the impact and application of the feedback. For nursing students to create knowledge and translate it to practice, they need to actively seek information, synthesise it and apply to subsequent tasks (Molloy et al., 2020). So, feedback is in essence a shared responsibility (Carless and Winstone, 2023; Winstone et al., 2021), that is, higher education teachers/markers need to equip students with strategies for taking productive action on feedback information, and students need to engage with and use the feedback provided (Carter et al., 2019). Shifting feedback to a learning-centred process is invaluable for nursing students so they can take up and use feedback to improve any subsequent work, whether theoretical or practical. de Kleijn (2021) supports the notion that student feedback literacy is related to the development of teacher feedback literacy.

Therefore, it is important that as part of the calibration (and moderation) practices undertaken with marking teams, that feedback literacy is addressed concurrently. A gap is currently present in the literature as these practices are not discussed together. This research explored nursing academics understanding of each and how they relate to each other. With these considerations and background, the aim of the study was to investigate and establish an understanding of nursing academic staff perceptions and practices in calibration processes and associated feedback literacy.

2. Method

2.1. Design

This research study used complementary theoretical underpinnings in participatory and person-centred research methodology. The methodology stemmed from a critical ontological perspective where multiple realties are accepted. This relates to the ontological perspective of the research study, that is, a person-centred perspective (McCormack et al., 2017), where the intent of the research was to explore the reality that exists within the context where the research was conducted. Person-centred research adheres to the belief that social reality is defined by persons (that is, it is not fixed), and is influenced through institutions and culture. This is revealed in forms of socially constructed knowledge, some of which are more valued than others. When applied to this research, where the emphasis was to better understand the research context from the experience of the participants, this makes it suitable to

meet the stated aim of the study. Since academic staff are actively engaged in the marking process, they are in the best position to provide feedback and insight into potential practice improvement.

2.2. Ethical considerations

The study was approved by the University xxx Human Research Ethics Committee (2022/002). All participants were provided with a detailed information sheet outlining the commitment required. Written consent was obtained before commencing interviews. Anonymity and confidentiality were assured, and all transcripts de-identified. All participants provided a pseudonym. Participants were able to withdraw from the study without impact prior to the group interview.

2.3. Setting, participants and data collection

Participants were recruited from staff in the School of Nursing at a regional university in Australia with six campuses. Inclusion criteria were any nursing academic staff (continuing, contract or sessional) who undertake marking of any form of summative assessment in undergraduate and/or postgraduate subjects in the School of Nursing at the University. This assessment was inclusive of essays, case studies, digital presentations, reflections, and other varieties of assessments submitted for grading.

The setting of is a School of Nursing in a regional university with six campuses offering the pre-registration Bachelor of Nursing program, and two of these offering postgraduate courses. Student numbers in the pre-registration program are approximately 600 in each of the three years, so large teams are required to manage marking loads. Postgraduate numbers are generally between 50 and 200 for a subject instance, so marking is required to be spread across teams of academic nurses, both permanent, contract and sessional. Common practice is for a team to meet prior to an assessment being due and discuss expectations and practices of marking. This is often accompanied by marking of three assessments independently and reviewing together to seek consensus. All sessional markers are paid for this meeting and marking in addition to the number of assessments marked. Payment is based on a formula aligned to word count (or equivalent).

Emails were sent to all academic nursing staff from an independent administrative professional staff member. Two follow up emails were sent to encourage participation. Data were collected from eight nursing academic staff via three focus group online interviews between June 9, 2022, and June 22, 2022. Focus group one consisted of two participants, a permanently appointed subject coordinator and a sessional tutor/marker. Focus group two comprised two participants, a subject coordinator on contract and a sessional tutor/marker. The final focus group

Table 1
Participant demographics.

Group interview	Participant pseudonym	Role	Years experience
1	Asha	Subject coordinator; permanent	3
1	Betty	Sessional tutor and marker	3
2	Clara	Subject coordinator; contract	1
2	Eve	Sessional tutor and marker	4
3	Kylie	Subject coordinator; permanent	5
3	Sally	Sessional tutor and marker	7
3	Chloe	Sessional tutor and marker	10
3	Bob	Subject coordinator; permanent	6

entailed four participants, two permanently appointed subject co-ordinators and two sessional tutor/markers (see Table 1 for details). As all researchers were academics in the School of Nursing, an independent facilitator conducted all interviews alone to ensure participants felt safe to speak freely and no power imbalances or bias were present. Key questions asked in the semi-structured group interviews were around marking practices and feedback processes the participants had been involved in, best practice in calibration processes, and what feedback literacy meant to the participant in terms of the process of calibration. Participants were encouraged to speak freely and provide as much detail as possible with accompanying examples from their calibration practices and/or experiences and associated feedback. With each participant's permission, interviews were audio recorded and field notes taken. Interviews lasted between 26 min and 59 min (mean 46 min). All interviews were independently transcribed verbatim. No repeat interviews were conducted, and no participants withdrew from the study.

2.4. Qualitative rigour

The study adhered to the principles of Lincoln and Guba's criteria of credibility, dependability, confirmability, and transferability (Lincoln and Guba, 1985). Credibility was established through an independent interviewer with education and facilitation expertise conducting all interviews. Participants ranged in experience and employment status, hence ensuring a range of data sources and experiences. Three researchers independently transcribed for accuracy. Six researchers independently, and then collaboratively, analysed and reached consensus. Using an external nursing education expert, an inquiry audit occurred to evaluate the quality of data, coding, and themes. In this way, dependability was assured. Confirmability was established through an audit trail where all interviews were audio recorded and field notes taken. Finally, transferability was ensured by the rich description of the findings that were related to previous studies.

This study employed a person-centred approach, underpinned by the ontological perspective that social reality is defined by persons and by essence is not fixed. This method aligns to Braun and Clarke's (2021, p. 206) notion that saturation is not consistent with all forms of qualitative inquiry and is an illogical approach to data, since "there are always new theoretical insights to be made as long as data continues to be collected and analysed". We acknowledge the sample size may be considered small, however the meaning interpreted through the participant voices has provided rich data addressing the research aim. Consistent findings occurred in all focus group interviews, but we recognise that new meanings may be found with more participants involved.

2.5. Data analysis

NVivo 12 software was used to manage the data. Braun and Clarke's (2022) six phases of reflexive thematic analysis were used for data analysis. This robust and rigorous analysis method can be used reliably to determine key themes from the data (Braun and Clarke, 2022).

A researcher from the team (CJ) listened independently to all recordings whilst reading the transcripts to ensure accuracy in transcription. Following this, five researchers (RM, KL, HP, KR, CA, SB) who did not have access to the audio recordings read the transcripts repeatedly to ensure familiarity with the data (phase 1). This process was done to ensure confidentiality of participants who the researchers may have known. The five researchers then broke into two groups and began the process of analysis, formulating meaning into codes and themes (phases 2 and 3). They then convened twice as a whole to discuss findings and reach consensus on themes (phase 4). The themes were discussed and vetted with all authors, and a nursing education expert external to the study, to ensure accuracy of themes (phase 5). Phase 6 followed with writing the findings into publication.

3. Results

The study collected information from eight nursing academics, seven females and one male. Five were either permanent or contract staff members and two were casual sessional staff.

Their experience ranged from one to ten years (mean experience four years) (Table 1).

From the data four overarching themes were identified: rubrics, calibration, feedback and justice. These themes were conceptualised as an umbrella depiction where the processes and practices associated with rubrics, calibration and feedback come together to ensure that students receive justice when their assessments are marked (see Fig. 1).

3.1. Rubrics

According to the participants rubrics are a pre-requisite for calibration. They conceived the purpose of rubrics as providing objectivity and clarity when marking with the underpinning value of ensuring equity. Eve described the necessity of this process using the analogy of a lens. In her words: "*But that the marks were going to be equitable and that we were all going to be, you know, using the same lens as we provided feedback*".

Most of the participants felt that for both students and the marking team, rubrics established the expectations for achieving specific outcomes. The participants perceived that rubrics enabled students to develop assessments with a clear understanding of the requirements across the various elements of an assessment. However, it was noted that the academic participants felt that students did not always understand the rubric or its purpose:

"If we set them up a little bit more about understanding what it is to read a rubric and understand a rubric and how it is applied to an assessment, I feel like that might help with feedback literacy too. So, we're thinking about having an activity in tutorial about understanding and marking with it. And yeah, getting them to go through an essay, like a provided essay, with a rubric"

(Asha).

Most participants regarded markers as having a reliance on the rubric to help them differentiate between various categories and grades (e.g. pass versus credit). This was only made possible however where descriptors were specific and objective. According to the participants, the rubric was also helpful where markers knowledge of the subject was limited, as was the case at times with casual sessional markers,

"So, the last [assessment] we did last semester was really succinct with what was needed to get a pass a credit, you know, distinction and so. For that reason, you can come back to the student with you know you have met this criteria, therefore, you've been given a pass or whatever, because they've fulfilled that. So I've always found the rubric is my, you know, Bible, to fully go by"

(Betty).

3.2. Calibration

The participants pointed to the process of calibration as enabling rubrics to be operationalised in teams, providing consistency of marking approaches across the marking team. Most described the calibration process as a two phased process. The first phase involved the subject coordinators identifying three or more assessments that represented different categories/grades. Markers were then expected to grade the anonymised assessments and return these to the subject coordinator. The second phase involved a discussion of these marks at the calibration meeting. Kylie shared her process of leading a calibration meeting:

"We'll go through the rubric and the findings like where the spread of results were and just have a look at you know, get everyone to have a
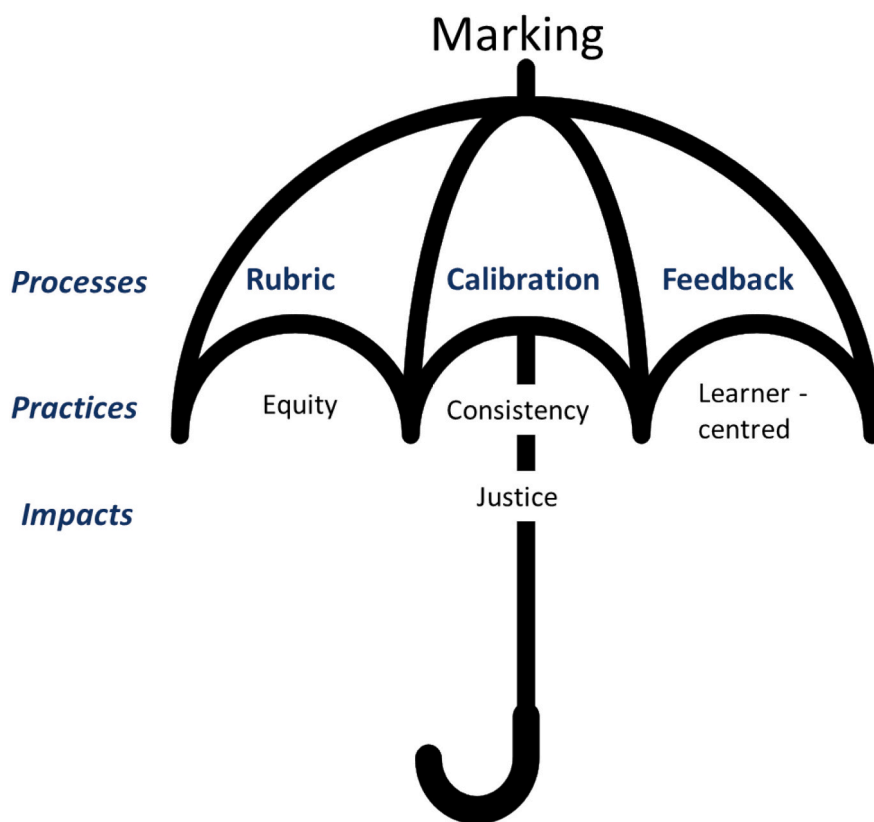
Marking

Fig. 1. Conceptual map.

little look at where we all sat and try and get some consensus around what we're looking for in each and objectively consider the criteria".

There was consensus among the participants that the calibration meeting was generally a one-hour recorded meeting, providing a record for markers who may not have been able to attend. This meeting was described as a rich discussion between markers, helping them understand how to apply the rubric, gain important perspectives on what the subject coordinator envisaged, and reach consensus over what constituted each category of achievement for this assessment. In Bob's words:

"I just get 3 assessments out and those assessments will be deidentified and all the people that are marking in the subject will receive those three papers and we all of us will mark the paper based on the rubric that we have built into the subject's assessment and then of all given our own little code. I just usually colour code everyone and then once everyone marked those three papers and provided feedback on them, they send it back to me and I then uploaded onto a rubric so that we can sort of see where everyone marked in each of the criteria, and then we once we do that, we do a critical discussion based on the question and the criterion"

(Bob).

Participants explained that this consistency was important so that students received fair or impartial marking, as Sally stated, "*yeah, actually I understand, and I get that and then that sets you off on the path when you go to mark all your papers. This is like well, yes, now I really know what I'm looking for*".

3.3. Feedback

Participants aimed to provide nuanced and constructive feedback that would help students improve their future performance. To achieve this, participants tried to provide learner centred feedback that was useful, specific, and kind. The ideal feedback reported by subject

coordinator participants was for all markers to address all elements of the rubric as well as an overall comment. Other sessional academic participants described developing their own standardised approach to feedback, including the use of quick marks and the feedback sandwich. They felt it was helpful to provide specific examples of where students could improve, such as where a student had not used the appropriate literature to support their case or incorrect referencing style. Participants agreed that it was important to be kind when providing feedback, especially where an assessment did not address required elements or an expected standard.

"So, when I've come up with a difficult paper it's helped me with a couple of things, but, trying to be constructive but positive in that feedback because you know, try and soften the blow. Look, you know you would have found your marks would have improved if you've if you've done blah blah blah"

(Clara).

There were concerns by participants that students understanding of feedback was limited because they tended to only look at the mark, and not read the comments within and at the end of assessment. The effect of this was that students would not consider the advice on areas where they could improve their performance moving forward in future assessments, "*So, I don't know how much students would read like my quick mark (including) … good point, or great point, or incorrect referencing*" (Chloe).

3.4. Justice

Participants described several downsides or threats to achieving equitable and consistent marking and providing learner-centred feedback, that is, to providing just practices for both student and marker. Time was woven intricately with Justice. This impacted the marker, particularly in terms of time efficiency, and the student in terms of feedback quality which could enhance or grow their feedback literacy.

Markers noted that considered feedback took time to provide, for example, Kylie stated that on top of reading the submission, "to give good and succinct feedback is probably 10 minutes in itself". All participants strongly agreed that the current allocated marking times for various word counts were not enough to provide students with equitable marking and constructive learner-centred feedback.

"I think particularly the thing that I find really hard is setting expectations for the markers, like we expect you to answer every section of the marking rubric, knowing full well that that is going to take more than 10 minutes. That for me is huge. Because we're already working in an industry where we're under resourced, undervalued, like I just feel like I'm just perpetuating that problem. And it really, really irks me to be honest. It is the biggest part of the whole process for me, that I wish I could change"

(Asha, subject coordinator).

Lack of time to provide considered feedback, when coupled with students seeming disengagement with feedback provided, focusing on a score only, markers felt despondent about the time taken to ensure students understood their result. Bob shared:

"And I, and that's tricky because like I put up announcements about that and stuff like that and we went through how to read your feedback in tutorial, right, but I mean it's hard. If we set them up a little bit more about understanding what it is to read a rubric and understand a rubric and how it is applied to an assessment, I feel like that might help with feedback literacy too. So, we're thinking about having an activity in tutorial about understanding and marking with it. And yeah, getting them to go through an essay, like a provided essay, with a rubric".

(Bob, subject coordinator)

Clara added that greater student understanding, and therefore a sense of justice related to the mark awarded, could be achieved through practices such as:

"And I think releasing like, general feedback. Like we always have put the average mark on there and I always will release it before we release all of the marks so there's some expectation about what the average mark is when they get their mark".

Contributing to the practice of justice for markers and students were aspects relating to assumed knowledge and skill. Participants who were subject coordinators had learnt to develop rubrics using experience and corporate knowledge rather than direct education. This limited their ability to develop objective actionable criteria, as Chloe stated, "You know you don't go into [academia] as an educator or whatever, you've got to start at that basic level again". In addition, this influenced marking practices. Participants described that often they were left to their own devices when it came to the processes of marking, with it being assumed they knew how to navigate the marking processes, including using electronic marking systems.

"I just had one of my colleagues come on board with XXX and she rang me up with her first lot of marking and she was in all sorts. So, I went through the whole thing with her and only that I had someone go through that with me; otherwise, I wouldn't have known what was expected"

(Betty, sessional tutor).

These instances led participants to agree that all staff involved in marking would benefit from formal training – from navigating the online systems, to applying rubrics and providing feedback. Bob (subject coordinator) described how "*better support*" for all markers would assist in building a culture of learning and feedback. This would lead to fair processes for both markers and students, and contribute toward

enhanced feedback literacy.

4. Summary

The collective goal for these participants was to ensure students received justice when their assessments were marked. That is, that the marking process was just, where assessments were marked against an objective criterion (rubric) that was applied consistently (calibration) and received marks that reflected a student's achievement under a specific element. Moreover, the rubric was available so that students understood what was required for an individual assessment. Lastly students were provided with constructive and specific feedback that they would use to improve their future performance. Hence the conceptualisation of the marking process as an umbrella (Fig. 1).

5. Discussion

The findings from this study have captured the perceptions of academic staff in relation the processes, practices and impact of calibration on practices in higher education marking, in relation to both staff and students. We have captured the significance of feedback literacy associated with marking calibration practices for nursing academics and students. Participants in our study provided a lens of the perceptions of staff engaged in marking within a pre-registration undergraduate nursing program, pointing to the staff's own experiences as well as their perceptions of how students experience this. We present these findings as a depiction of an umbrella framework with the processes being the fabric (rubric, calibration, and feedback), the practices the spokes (equity, consistency and learner-centred), and the impact as the handle or central support (justice) (Fig. 1). As with the structure of an umbrella, all elements are required for effectual purpose.

This research points to marking as a skill that takes time to develop and highlights the importance in developing both marking experience and team support to achieve consistency in marking. The participant's understanding that the process of calibration improves staff consensus, provided a sense that the process of approaching marking with the same mind as colleagues, translated to better practices and ultimately the outcome of consistency and justice. This aligns with the notion outlined by Herbert et al. (2014) that calibration informs a common understanding in markers so that consistency in judging divergent responses to assessment tasks occurs. Mason and Roberts (2023) and Ragupathi and Lee (2020) also found that calibration enables collegial conversations which build confidence and consistency in marking teams. In addition, this process enables markers to be clear about the required level of feedback. Commonly, markers provide surface level feedback with focus on writing mechanics (Dressler et al., 2019) and without specificity or clear suggestion of expectation(s) which are required for deeper processing, leading to future learning and improvement (Reynders et al., 2020; Baranczyk and Best, 2020). Providing consistent and meaningful feedback facilitates learners' ability to recognise and enhance future work (Börjeson and Carlsson, 2020), thereby increasing feedback literacy.

While the participants in this study welcomed the use of rubrics and the support of a team approach, there was a sense that novice markers and academics are still developing these skills. Care needs to be taken to avoid power imbalance so that all markers, irrespective of experience, feel safe to clarify and ask critical questions (Mason and Roberts, 2023). Research however would indicate that while experience may improve confidence, it does not necessarily improve consistency (Bloxham et al., 2016). Experience can create habits and ways of marking that could prove to be a barrier to consistency. It may even deter experienced markers from being involved in calibration if it is seen to differ from the processes they have established for themselves. Therefore, focusing on developing markers' shared understanding of assessment requirements and using calibration to develop and maintain standards is essential (Mason and Roberts, 2023). Sadler (2013) outlines this social process of

calibration as requiring synergistic relationships in the marking team, enabling consensus to be reached on mark allocations for an assessment.

The ability to provide consistency in marking and grades is the hallmark of equitable assessment, however it is not a simple task. According to O'Connell et al. (2016): "Academics attach qualities to students' work according to their varying professional knowledge and experience" (p. 336). Bloxham et al. (2016) and Grainger et al. (2016) concur, noting that tacit knowledge and associated expectations vary between markers. These, alongside potential biases and/or prior knowledge of students may increase variation in marking (Orr and Bloxham, 2013). This is compounded by larger numbers of markers, particularly sessional academics, in the marking team (Crawford and Germov, 2015). Through consensus building and a supportive open environment, shared understanding of assessments and the role of the marker can be explored and affirmed, building consistency and confidence (Grainger et al., 2019). Our research demonstrates the importance of markers having core values related to justice and consistency. This was evident in the participants' commitment to calibration processes and the consideration that influenced their decision making around the allocation of marks and the type of feedback they provide.

Time is an essential element that needs to be addressed if marking consistency is to be achieved. Time impacts both the process of marking and the ability to provide thoughtful and meaningful feedback to students. In higher education the time spent marking and providing feedback is challenged by competing priorities of teaching and research as well as the reliance on sessional staff who are paid according to the word count of an assessment. The latter was a significant concern for participants in this present study. The findings of this study point to the importance of processes such as well-developed rubrics, calibration processes and the provision of feedback as being central to consistency however each of these steps requires a commitment of time. This is a key consideration for workload allocation, particularly when advocating for collaborative and collegial processes of calibration practices (Richardson et al., 2019).

It has been conventionally assumed that having knowledge of subject matter is sufficient for feedback to be conducted well, but this study shows this is not always the case and that substantially increased levels of pedagogic competence are needed in addition to assist student learning (Boud and Dawson, 2023). There is no 'magic formula' for feedback (Sadler, 2009), and as such, no amount of simplistic tips and tricks or training in specific practices will be guaranteed to improve feedback. What is needed instead is an increase in overall competence in teachers' feedback literacy (Carless and Winstone, 2023). Boud and Dawson (2023) agree, stating that there are many complex factors associated with this, particularly in relation to navigating the differences between assessment and feedback, and the purpose of feedback in facilitating future learning and work.

Carless and Boud (2018) outline four features of student feedback literacy: "appreciating feedback processes; developing capacities in making judgments; managing affect; and taking action to use feedback" (p. 1323). Our findings build on this by adding a sense of equity, justice and learner-centred practices as underpinning elements to the student feedback literacy experience. Across the interviews, there was strong sense of the importance for staff practices of these elements in order for fairness to be achieved.

Providing opportunities for students to develop feedback literacy requires processes that support learner-centred marking. The emphasis from teachers teaching to learners learning is consistent with transformative pedagogies in learning and teaching (Van Schalkwyk et al., 2019).

Based on this premise, the authors recommend nursing curricula are structured to enable ongoing opportunities for students to be immersed in feedback following assessments, affording time to process and enact the feedback. This is only possible if markers and teaching staff are clear on feedback literacy (Maleck et al., 2022). Therefore, it is important to design nursing curricula with emphasis on developing feedback literacy,

building in formative and summative assessment tasks where students engage in learning that facilitates the development of course learning outcomes. This is critical in first year units so that students can be brought to a minimum level of feedback literacy (Maleck et al., 2022). Feedback literate nursing educators must develop mindsets that understand the competing functions of feedback and use practices such as calibration to enable practices that have potential to enrich student learning (Carless and Winstone, 2023).

5.1. Limitations

Although there was a semi-structured script and the interviewer was not known to the participants, researchers bring their own biases and ways of being to the data collection process, and this may have impacted the consistency of the data collection process. Another limitation is the number of participants in this study, which may be considered small despite the aligning of responses and themes. The sample was a convenience sample of permanent, contract and sessional academics voices from one School of Nursing in Australia (across six campuses). While participants came from a variety of campuses across regional and metropolitan areas, the sample is not necessarily representative of the wider population and represents participants' perceptions at a single point in time.

This project took into consideration the perceptions of staff involved in marking. What it does not capture is the voices of the students. To fully understand feedback literacy as a student experience, further research involving students would be advantageous.

6. Conclusion

The benefits of participating in the project are to inform the body of knowledge around academic practices in calibration and the connection to feedback literacy. Input could potentially help to influence educational strategy and policy at a School level and possibly Faculty and institution. Involvement may help improve both novice and experienced academics own scholarship, learning, teaching and academic practices. Ongoing development and collaborative support may achieve consistency within the calibration process. This project highlights the need for enhanced feedback literacy that integrates principles of equity, justice, and learner-centeredness. Fundamentally calibration guides educators toward holistic approaches that foster consistency, equity, and thorough feedback practices.

CRediT authorship contribution statement

Rebekkah Middleton: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. **Kelly Lewer:** Formal analysis, Writing – original draft. **Carolyn Antoniou:** Formal analysis, Writing – original draft. **Helen Pratt:** Formal analysis, Writing – review & editing. **Suzanne Bowdler:** Formal analysis, Writing – review & editing. **Carley Jans:** Writing – review & editing. **Kaye Rolls:** Formal analysis, Methodology, Writing – original draft, Writing – review & editing.

Declaration of competing interest

All authors work or have worked in the School of Nursing where data was collected. No authors were involved in data collection or had any knowledge of the participants. An independent facilitator collected all consent forms and conducted all group interviews. No authors have any conflicts of interest related to this publication.

the interviews and the independent facilitator who conducted the interviews.

Funding

No funding has been received to support this study.

The authors declare no use of generative AI was used in the writing of this manuscript.

References

Baranczyk, M.C., Best, C.A., 2020. How does instructor feedback help and hurt students' academic performance? Scholarship of Teaching and Leanirng in Psychology 6 (1), 1–14. https://doi.org/10.1037/stl0000169.

Bloxham, S., Price, M., 2013. External Examining: Fit for Purpose? Stud. High. Educ. 1 (13), 1–17.

Bloxham, S., den Outer, B., Hudson, J., Price, M.E., 2015. Let's stop the pretence of consistent marking: exploring the multiple limitations of assessment criteria. Assess. Eval. High. Educ. 41 (3), 1–16.

Bloxham, S., den Outer, B., Hudson, J., Price, M.E., 2016. Let's stop the pretence of consistent marking: exploring the multiple limitations of assessment criteria. Assess. Eval. High. Educ. 41 (3), 466–481.

Börjeson, F.C., Carlsson, C.J., 2020. Talking about writing – designing and establishing writing feedback and turorials to promote student engagement and learning. Journal of Academic Writing 10 (1), 128–135. https://doi.org/10.18552/joaw.v10i1.604.

Boud, D., Dawson, P., 2023. What feedback literate teachers do: an empirically-derived competency framework. Assess. Eval. High. Educ. 48 (2), 158–171. https://doi.org/10.1080/02602938.2021.1910928.

Boud, D., Lawson, R., Thompson, D.G., 2015. The calibration of student judgement through self-assessment: disruptive effects of assessment patterns. High. Educ. Res. Dev. 34 (1), 45–59.

Braun, V., Clarke, V., 2021. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qualitative Research in Sport, Exercise and Health 13 (2), 201–216. https://doi.org/10.1080/2159676X.2019.1704846.

Braun, V., Clarke, V., 2022. Thematic analysis: A practical guide. Sage.

Carless, D., Boud, D., 2018. The development of student feedback literacy: enabling uptake of feedback. Assess. Eval. High. Educ. 43, 1315–1325.

Carless, D., Winstone, N., 2023. Teacher feedback literacy and its interplay with student feedback literacy. Teach. High. Educ. 1–14 https://doi.org/10.1080/13562517.2020.1782372.

Carter, R., Halcomb, E., Ramjan, L.M., Wilson, N.J., Glew, P., Salamonson, Y., 2019. Does the use of annotated exemplars by nursing students predict academic performance? A cohort study. Nurse Educ. Today 80, 34–39.

Cleary, T.J., Konopasky, A., La Rochelle, J.S., Neubaser, B.E., Durning, S.J., Artino, A.R., 2019. First-year medical students' calibration bias and accuracy across clinical reasoning activities. Adv. Health Sci. Educ. 24, 767–781.

Crawford, T., Germov, J., 2015. Using workforce strategy to address academic casualisation: a University of Newcastle case study. J. High. Educ. Policy Manag. 37 (5), 534–544. https://doi.org/10.1080/1360080X.2015.1079394.

de Kleijn, R.A.M., 2021. Supporting student and teacher feedback literacy: an instructional model for student feedback processes. Assess. Eval. High. Educ. https://doi.org/10.1080/02602938.2021.1967283.

Dressler, R., Chu, M., Crossman, K., Hilman, B., 2019. Quantity and quality of uptake: examining surface and meaning-level feedback provided by peers and an instructor in a graduate research course. Assess. Writ. 39, 14–24. https://doi.org/10.1016/j.asw.2018.11.001.

Grainger, P., Adie, L., Weir, K., 2016. Quality assurance of assessment and moderation discourses involving sessional staff. Assess. Eval. High. Educ. 41 (4), 548–559.

Grainger, P., Crimmins, G., Burton, K., 2019. Assuring the quality of curriculum, pedagogy and assessment across satellite campuses. J. Furth. High. Educ. 43 (5), 589–600. https://doi.org/10.1080/0309877X.2017.1386286.

Herbert, I.P., Joyce, J., Hassall, T., 2014. Assessment in higher education: the potential for a community of practice to improve inter-marker reliability. Acc. Educ. 23 (6), 542–561.

Lincoln, Y.S., Guba, E.G., 1985. Naturalistic Inquiry. Sage Publications, Newbury Park, CA.

Maleck, B., Boud, D., Carless, D., 2022. Eliciting, processing and enacting feedback: mechanisms forembedding student feedback literacy within the curriculum. Teach. High. Educ. 27 (7), 908–922. https://doi.org/10.1080/13562517.2020.1754784.

Mason, J., Roberts, L.D., 2023. Consensus moderation: the voices of expert academics. Assess. Eval. High. Educ. https://doi.org/10.1080/02602938.2022.2161999.

McCormack, B., van Dulmen, S., Eide, H., Skovdahl, K., Eide, T., 2017. Person-centred Healthcare Research. Wiley-Blackwell.

Molloy, E., Boud, D., Henderson, M., 2020. Developing a learning-centred framework for feedback literacy. Assess. Eval. High. Educ. 45 (4), 527–540.

O'Connell, B., De Lange, P., Freeman, M., Hancock, P., Abraham, A., Howieson, B., Watty, K., 2016. Does calibration reduce variability in the assessment of accounting learning outcomes? Assess. Eval. High. Educ. 41 (3), 331–349.

Orr, S., Bloxham, S., 2013. Making judgements about students making work: lecturers' assessment practices in art and design. Arts Humanit. High. Educ. 12 (2–3), 234–253. https://doi.org/10.1177/1474022212467605.

Quigley, D., 2021. When I say… feedback literacy. Med. Educ. 55, 1121–1122.

Ragupathi, K., Lee, A., 2020. Beyond fairness and consistency in grading: the role of rubrics in higher education. In: Sanger, C., Gleason, N. (Eds.), Diversity and Inclusion in Global Higher Education. Palgrave Macmillan. https://doi.org/10.1007/978-981-15-1628-3_3.

Read, B., Francis, B., Robson, J., 2005. Gender, bias, assessment and feedback: analyzing the written assessment of undergraduate history essays. Assess. Eval. High. Educ. 30 (3), 241–260.

Reynders, G., Lantz, J., Ruder, S.R., Stanford, C.L., Cole, R.S., 2020. Rubrics to assess critical thinking and information processing in undergraduate STEM courses. International Journal of STEM Education 7, 9. https://doi.org/10.1186/s40594-020-00208-5.

Richardson, J., Wardale, D., Lord, L., 2019. The 'double-edged sword' of a sessional academic career. High. Educ. Res. Dev. 38 (3), 623–637. https://doi.org/10.1080/07294360.2018.1545749.

Rust, C., 2009. Assessment standards: a potential role for subject networks. J. Hosp. Leis. Sport Tour. Educ. 8 (1), 124–128.

Sadler, D.R., 2009. Indeterminacy in the use of preset criteria for assessment and grading. Assess. Eval. High. Educ. 34 (2), 159–179.

Sadler, D.R., 2013. Assuring academic achievement standards: from moderation to calibration. Assessment in Education: Principles, Policy & Practice 20 (1), 5–19.

Van Schalkwyk, S.C., Hafler, J., Brewer, T.F., Maley, M.A., Margolis, C., McNamee, L., Meyer, I., Peulso, M.J., Schmutz, A.M.S., Spak, J.M., Davies, D., 2019. Transformative learning as pedagogy for the health professions: a scoping review. Medical Education in Review 53 (6), 547–558. https://doi.org/10.1111/medu.13804.

Winstone, N., Pitt, E., Nash, R., 2021. Educators' perceptions of responsibility-sharing in feedback processes. Assess. Eval. High. Educ. 1–14. https://doi.org/10.1080/02602938.2020.1748569.