

Undergraduate Science and Agriculture Students Perceive Improved Academic Judgement Upon Use of an Assessment Literacy Module

Melissa J. Saligari^a, Jennifer L. Fox^a, Claudia A. Rivera Munoz^b, Sarah Frankland^c, and Amber Willems-Jones^d

Corresponding author: Melissa J. Saligari (mel.saligari@unimelb.edu.au)

^aSchool of BioSciences, Faculty of Science, The University of Melbourne, Parkville VIC 3010, Australia;

^bMelbourne Centre for the Study of Higher Education, Faculty of Education, The University of Melbourne, Parkville VIC 3010, Australia

^cSchool of Agriculture, Food and Ecosystem Sciences, Faculty of Science, The University of Melbourne, Parkville VIC 3010, Australia;

^dDepartment of Biochemistry and Pharmacology, School of Biomedical Sciences, Faculty of Medicine, Dentistry and Health Sciences, The University of Melbourne, Parkville VIC 3010, Australia.

Keywords: assessment literacy, academic judgement, student perceptions, staff perceptions

Abstract

Success in higher education is associated with both students' capacity for self-regulation and the ability to understand assessment processes. The Assessment Literacy Module (ALM) is an interactive online tool designed to promote development of students' own academic judgement and assessment literacy. Students act as assessors of work samples; they use the assessment rubric and are guided by expert marker feedback that relates the marking standards to the assessment outcome. In this study, we evaluated perceptions of ALM impact on the student learning experience by surveying staff ($N = 18$) and students ($N = 416$) involved in 18 undergraduate subjects. Students reported that after completing the ALM they had a better understanding of the assessment criteria (94%); they felt the module helped them prepare for their assessments (90%) and were more confident about their assessment quality (86%). Staff perceptions of the benefits of the ALM closely aligned with those reported by students. We recommend this tool for enhancing assessment literacy and developing academic judgement in undergraduate students of all levels. The ALM is of greatest utility in large cohort subjects as a sustainable approach for providing all students access to expert feedback. Future research will explore how the ALM benefits teaching practices.

Introduction

In undergraduate higher education, students often find the marking criteria of assessment tasks difficult to understand (O'Donovan, Price, & Rust, 2008), perceiving the descriptions as vague, with a lack of transparency about the spread of marks and how performance levels are evaluated (Chan & Ho, 2019). When assessment tasks also require academic judgement (being able to self-evaluate one's own work, and the work of others), students' difficulty in understanding can be compounded. Although assessments are often accompanied by a marking rubric, many students do not engage with or use it to evaluate their own work (Andrade & Du, 2005). In isolation, even the best-designed marking rubrics generally lack sufficient internal explanation on how to evaluate criteria, and so are at risk of poor implementation without an accompanying dialogue, guidance, or moderation process in place (Panadero & Jonsson, 2020).

In recent years educators have made greater attempts to describe and define the concept of student assessment literacy so that it can be explicitly taught to students (Hovardas, Tsivitanidou, & Zacharia, 2014; Luo & Chan, 2023; Smith, Worsfold, Davies, Fisher, & McPhail, 2013). Most recently, this multifaceted concept has been defined as:

'Students' perception, attitude and contextualised knowledge and skills in developing strategies to actively engage in assessment, monitor their learning, engage in reflective practice, and develop effective skills, to improve their learning and performance outcome.' (Hannigan, Alonzo, & Oo, 2022, p. 493)

An important aspect of the student learning journey is the development of self-regulated learning behaviours. Self-regulation is described as the ways students regulate their thinking, motivation, and behaviours associated with learning (Pintrich & Zusho, 2002). Self-regulated learners tend to be more effective learners and higher achievers than those who do not self-regulate (Pintrich, 1995). However, there are aspects of learning tasks that students do not have control over, such as marking rubrics. For students to be effective self-regulators, they need to be able to understand and apply marking rubrics and task instructions, and it is here that self-regulation and assessment literacy overlap.

The need for a pedagogical tool that can aid in teaching assessment literacy skills has been well articulated in the literature (Hendry & Anderson, 2013; Sadler, 2002; Smith et al., 2013; Tai, Ajjawi, Boud, Dawson, & Panadero, 2018), and the Assessment Literacy Module (ALM) used in this study has been shown to be an effective tool (Willems-Jones et al., 2023). The ALM was developed by the Melbourne School of Psychological Sciences at The University of Melbourne (UoM) to improve student assessment literacy at scale. The ALM has been implemented broadly across UoM subjects including in Agriculture, BioSciences, and Biomedical Sciences since 2021.

Ultimately, the ALM is pitched towards improving student understanding of assessment requirements alongside marking rubric criteria standards, so that students are better equipped to evaluate both others and their own work. The ALM is a Canvas-integrated interactive program designed to explicitly teach students assessment literacy contextualised within individual subjects. The module can be used synchronously or asynchronously, depending on the academic's needs. As a self-directed learning activity, the ALM provides a time efficient approach to develop student assessment literacy skills in large cohort subjects. In the module, students sequentially mark example assessments and provide feedback on each assessable criterion (Figure 1).

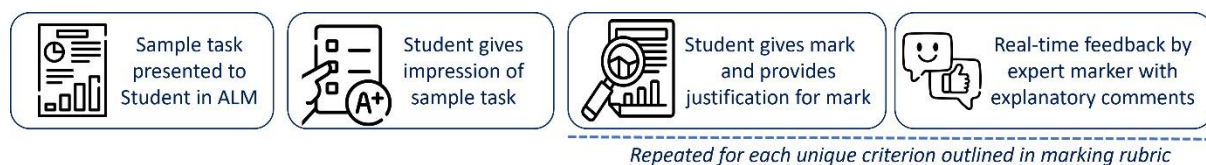


Figure 1. Activities completed by the student within the Assessment Literacy Module (icons created by Freepik, Flaticon).

Students step through each criterion of the example assessment and critically assess the work (that is, the work of others), providing a mark and justification for each criterion. The associated marking rubric is integrated within the program and students receive real-time pre-populated feedback illustrating how *their* assigned grade compares to the expert marker, with additional comments on why that specific grade was given by the marker. The process enables calibration

of the student's interpretation of the marking rubric and is the essence of how students learn and develop judgement (Boud, Lawson, & Thompson, 2013). Completion of the module takes approximately 1 hour and is often incorporated into a scheduled class. The examples are associated with summative assessments in the subjects deploying the ALM and may be authentic (de-identified) student work or custom-built by the academic.

This study evaluates staff and student perceptions on ALM use, expanding on our pilot study (Willems-Jones et al., 2023) through further staff and student survey responses. We evaluated how students use the ALM, whether it improves their confidence in assessment, whether it alters their rubric use, or improves their assessment preparation, potentially by fostering self-reflection. The research hypothesis – upheld in our pilot study – is that following ALM completion, students will perceive improvement in both their academic judgement and assessment literacy. The research questions underpinning the study include: (RQ1) Do staff feel that the ALM is an effective pedagogical tool for improving students' engagement with assessment and feedback practices? (RQ2) Do staff feel that ALM implementation improves the efficiency of their teaching practices? (RQ3) Do students perceive that ALM use has improved their academic judgement and assessment literacy? And (RQ4) do students perceive benefit in their assessments from ALM completion?

Methods

We evaluated perceptions of ALM impact on the student learning experience by surveying staff ($N = 18$) and students ($N = 416$) from 18 undergraduate subjects. Of the 18 subjects, student participants were enrolled in 10 subjects, staff participants taught into 14 subjects, and seven subjects had both staff and student study participants (AGRI10046 & 10049, BCMB20005, BIOM30001 & 30002, BIOL10008, GENE20001).

Staff Participants

Fifty academic staff (mostly subject coordinators) were invited to complete the survey, with 18 responding – almost meeting the target sample size of 20 participants (Malterud, Siersma, & Guassora, 2016). Academic staff were from multiple Schools (Agriculture and Food, Biomedical Sciences, BioSciences, Psychological Sciences, Population Health), teaching undergraduate subjects. Of these, 10 subjects were core to the degree/major, two were capstone experiences, one was an elective subject, and one was a university breadth subject. The student cohort size in these subjects varied from small (<30 students) to large (1200 students). Two staff participants each taught in two individual subjects. The study was approved by University of Melbourne Human Research Ethics Committee (HREC) (ID: 2022-23388).

Student Participants

A self-report survey was distributed to 3,000 students via Canvas announcements in 10 subjects across three undergraduate subjects (HREC ID: 2022-23773). Fourteen percent of students responded ($N = 416$). Table 1 summarises survey respondents across individual subjects. Survey data was accessed by researchers at least two weeks after grade certification, and all student data was de-identified before analysis.

Table 1. Student participants and their corresponding subjects that utilised the ALM during 2022-3 (N = 416).

Subject Name Code (Assessment type)	Year Level	N	% of cohort
Introductory Biology: Life's Machinery BIOL10008 (Practical Report)	1	6	1
Introductory Biology: Life's Complexity BIOL10010 (Practical Report)	1	16 94	4 27
Techniques Molecular Science BCMB20005 (Laboratory Report)	2	5 14	3 8
Cellular Metabolism & Disease BCMB30011 (Essay)	3	4	6
Frontiers in Biomedicine BIOM30001 (Peer Assessment)	3	24	5
Molecule to Malady BCMB30002 (Group Poster)	3	10	2
Thinking Scientifically SCIE20001 (Blog post, Article analysis)	2	26 121	4 14
Foundations of Agricultural Science 2 AGRI10046 (Group Lab Report)	1	96 ^a	90
Animal Production Systems AGRI10049 (Farm Report)	1	96 ^a	81
Genetics for Agriculture AGRI10051 (Exam Short-answer Questions)	1	96 ^a	84

^a Some students were enrolled in multiple ALM-containing subjects at the same time as core subjects in their degree. A single student responded to the survey twice and provided different responses. Both responses were included in analyses.

The subjects represented by the staff survey only partially overlap with the subjects in the student survey. This inconsistency was due to HREC timelines or because the student cohort had already been surveyed in our earlier pilot study. Four additional subjects were available for the student survey, but the relevant staff declined to be surveyed or were unavailable during the data collection window.

Survey and data analysis

Qualtrics was used to deploy staff and student surveys, with consent sought at the commencement of survey. Survey responses are mostly presented as figures, percentage response (agree/neutral/disagree) to Likert-scale questions, along with representative quotations drawn from free-text items to support and explain key trends. All participant responses were anonymised and aggregated by frequency and percentage of agreement/disagreement for Likert-scale responses across both surveys.

Analysis of the staff survey was based on three themes, 'Using the ALM', 'Perceptions of ALM', 'Reflections and future use', across a total of 31 items. The first and last of these sections comprised selected-choice and open-ended questions to allow for semi-structured response collection. The middle section comprised 5-point Likert scale agreement/disagreement items, with additional multiple-choice and open-ended questions. Analysis on both the quantitative (Likert scale) and qualitative (content) data was performed.

The content analyses involved two steps: firstly, survey responses were coded inductively into categories, and then these categories were combined into themes. In some instances, single staff responses were related to more than one theme, so total frequencies of themes are higher than total numbers of participants.

Analysis of the student survey was based on three themes, 'Using the ALM', 'Academic judgement', 'Future use of ALM'. Most of the survey items (total of 28) were 3- and 5-point Likert scale questions where participants ranked their level of agreement/disagreement with given statements, plus options for free-text. Both quantitative (for Likert items) and qualitative content analyses were carried out to identify themes across de-identified students' responses overall. Where comparisons were made of student grades in ALM-associated assessment tasks (Supplementary material, Table S1), a Student's *t*-test (1-tail distribution, unequal variances) was performed between survey participant grade and the broader cohort grades for the same assessment item.

Results

Staff perceptions and reflections on ALM use and student outcomes (RQ1).

Eighteen academic staff used the ALM to prepare students for at least one summative assessment in their subject. The number of examples presented in the ALM varied across the deployed subjects. Thirteen staff (72%) provided students with two example assessments, three staff provided three examples, and a further two staff provided more than three examples. Most of the staff participants ($n = 14$, 78%) indicated that the ALM was used as assessment literacy training for students, with completion being optional, however in one subject (Table 1, BIOM30001), module completion contributed to 5% of the subject grade. In free-text responses, the remaining three staff (17%) indicated that the ALM was a subject hurdle requirement.

The most frequently indicated purpose of the ALM by staff respondents was to help students understand the marking criteria and assessment task ($n = 12$, 67%). Other reported purposes included: preparation for assessment ($n = 10$, 56%), development of academic judgement to evaluate their own work and that of others ($n = 5$, 28%), or helping tutors understand feedback expectations ($n = 1$, 6%). Most staff participants perceived that students had a better understanding of assessment expectations and the assessment rubric after completing the ALM (Figure 2). Though some staff participants ($n = 4$, 22%) felt that it was difficult to encourage students to complete the ALM when it was not compulsory.

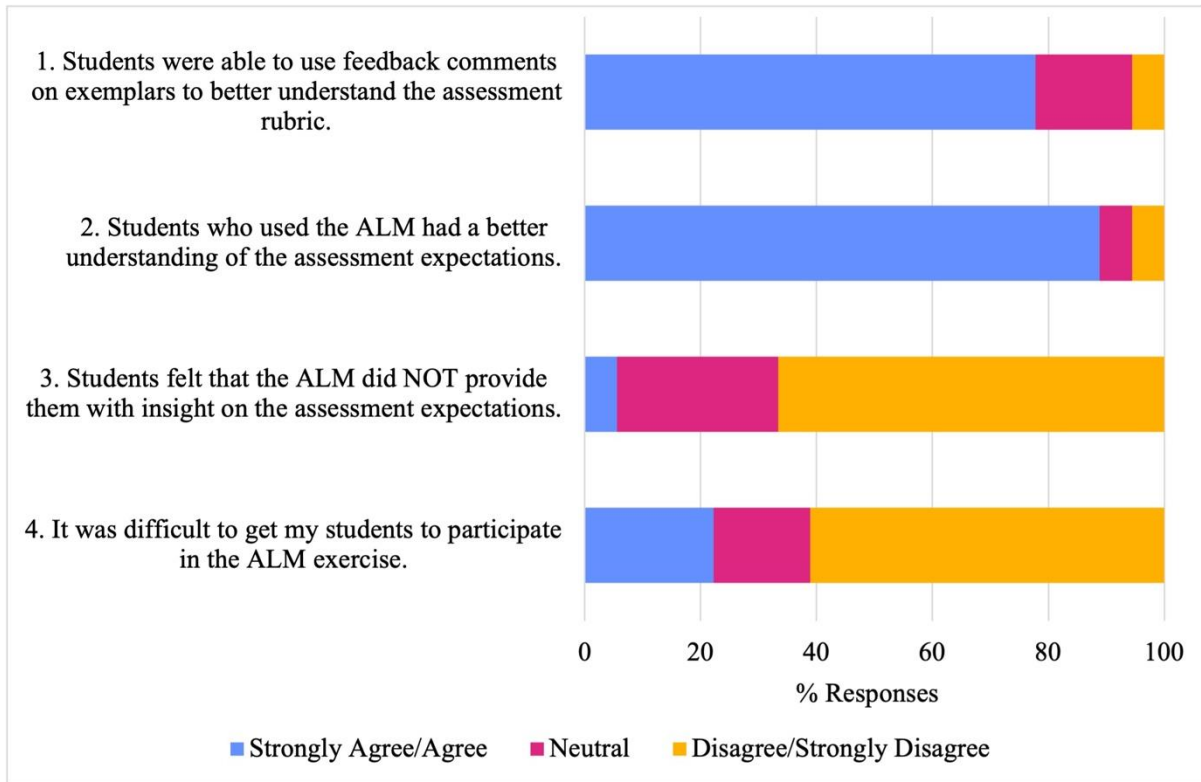


Figure 2. Staff perceptions of student-focused outcomes due to implementation of ALM in their subject ($N = 18$).

Most staff ($n = 14$, 78%) felt that students were able to use the ALM exemplars to better understand the assessment rubric. However, four staff indicated the opposite. Two staff clarified that this was because grades in the associated assessment did not increase over the previous year or that students did not engage meaningfully with the ALM.

Staff perceptions and reflections on administrative load and ALM implementation (RQ2).

Staff were asked to reflect on the use of the ALM in their subjects and respond to four statements related to the administration associated with implementing the ALM (Figure 3). Two common themes to emerge from Likert-scale based and open-response questions include the benefits to teaching practice (Figure 3, Statements 1 & 2), but also the increased workload due to the preparation required to implement the ALM (Figure 3, Statements 3 & 4).

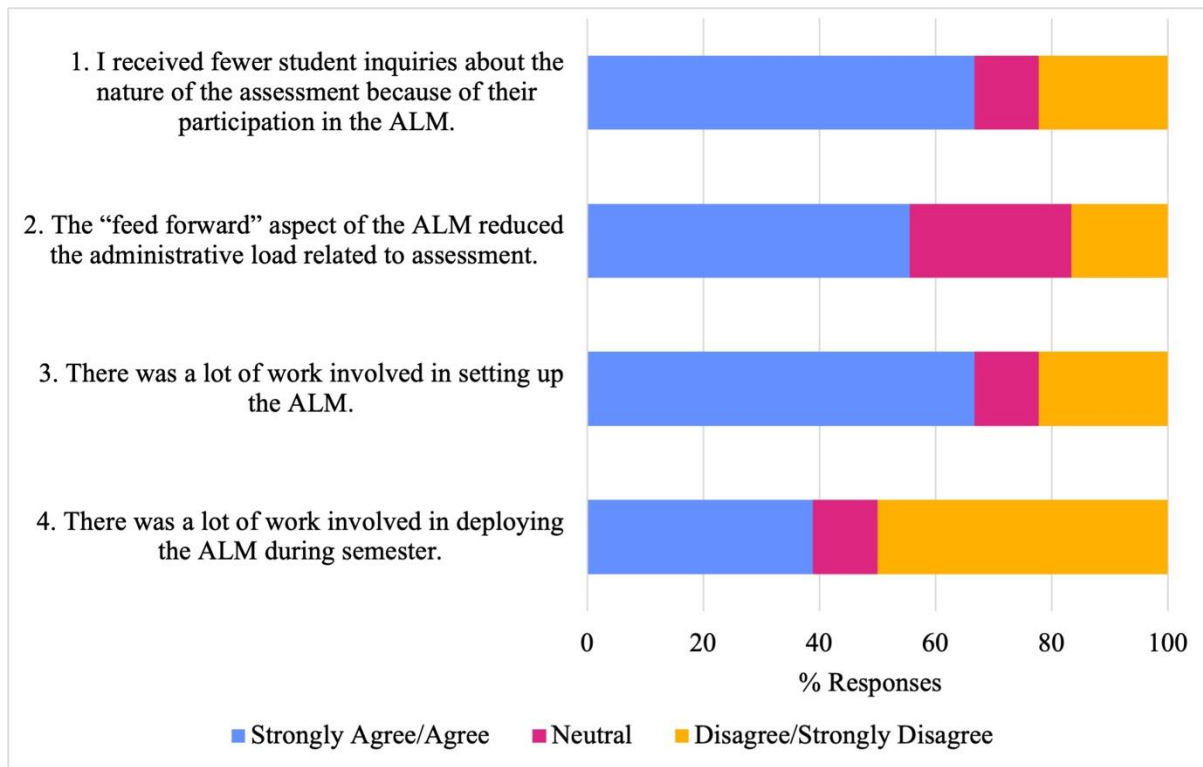


Figure 3. Staff perceptions of administrative load when implementing the ALM in their subject ($N = 18$).

Staff participants were asked to identify positive or negative aspects associated with using the ALM in open-ended responses. The most frequently reported positive aspects of ALM use included: (a) improvements to students’ understanding of the marking criteria and an increased confidence to developing their own assessments, reported by 13 staff, and expressed by S10: *“Makes students understand marking criteria better, relieves anxiety in some students as they see what essays get a good mark from experts”* (b) greater encouragement of students to review their own work, reported by six staff and explained by S5: *“Sets expectations for students, allows student to self-review, generates good educational outcomes overall”* (c) more instructive and immediate feedback for students, reported by three staff, and conveyed by S7: *“The feedback provided helps them identify the strengths and weaknesses of the sample reports, as well as how improvements could be made”* and (d) three staff felt that students’ agency was fostered by using the ALM to better understand how assessments are marked, expressed by S11: *“...it gave students the opportunity to be in the ‘driver’s seat’ of marking”*. Two staff (11%) identified an unintended positive consequence that the ALM was useful for the tutors’ marking process to reduce variability between markers: *“...training staff helped reduce inter-rater variability (decreased admin load).”* (S16)

The most common negative aspect of ALM use reported by staff in the open-ended questions was that it was time-consuming to set up the rubrics and examples ($n = 9$, 50%), while the proportion of all staff who found there was ‘a lot of work in setting up the ALM’ was closer to 65% ($n = 12$, Figure 3). Other common themes identified by staff were related to engagement with the module, whereby staff ($n = 4$, 22%) felt that students did not engage with the ALM in a meaningful way when it was a compulsory hurdle, for example: *“We*

can't ensure that students engage with it fully as they can just click through to tick off the hurdle.” (S7)

In addition, five staff (27.7%) felt their students tended not to value the ALM and did not complete it when it was an *optional* activity, as indicated by S18: *“Not all students saw the value in undertaking the activity”*. Two staff (11%) reported that they felt students had completed the ALM in multiple subjects and were “tired” and/or “disengaged” with it. However, this point of view does not align with student survey responses ($n = 271$, 89%) that indicated students would like the ALM implemented in other subjects (Figure 5, Statement 4 and Table 2, Theme 5). In terms of unintended adverse consequences of using the ALM, several staff ($n = 3$, 17%) identified the plagiarism of assessment examples, as well as the additional administrative work for staff when the task was classified as a hurdle, as expressed by one staff responder, S4: *“It was one further hurdle we needed to keep track of and to contact students about towards the end of semester. It is more work but not a huge amount more.”*

Student perceptions of assessment literacy and academic judgement resulting from use of the ALM (RQ3).

Of the students who completed the survey, 40% ($n = 167$) reported using the ALM in only one subject, while 45% indicated they used it in other subjects (15% did not specify). In terms of time spent using the ALM, a third of students (33%) spent between 30 minutes and one hour, including reading the examples; while 26% spent less than 30 minutes, and 20% spent more than one hour. Students were asked whether they used additional resources other than the ALM to complete the associated assessment in their subject. A total of 338 students responded, with almost all (96%) stating they used the marking rubric, followed by 64% who used the instruction document, a further 54% of respondents indicated they used the information from lecture videos to complete associated assessments.

Most students reported that the ALM was helpful in preparing their assessments, that they also perceived they had a better understanding of the criteria for the subject assessment, and they felt more confident about the quality of their assessments after completing the ALM (Figure 4), captured by a student reflection: *“[The ALM] showed me what was relevant and irrelevant. What made for a concise aim, sentences. What made for an adequate report and what was lacking.”*

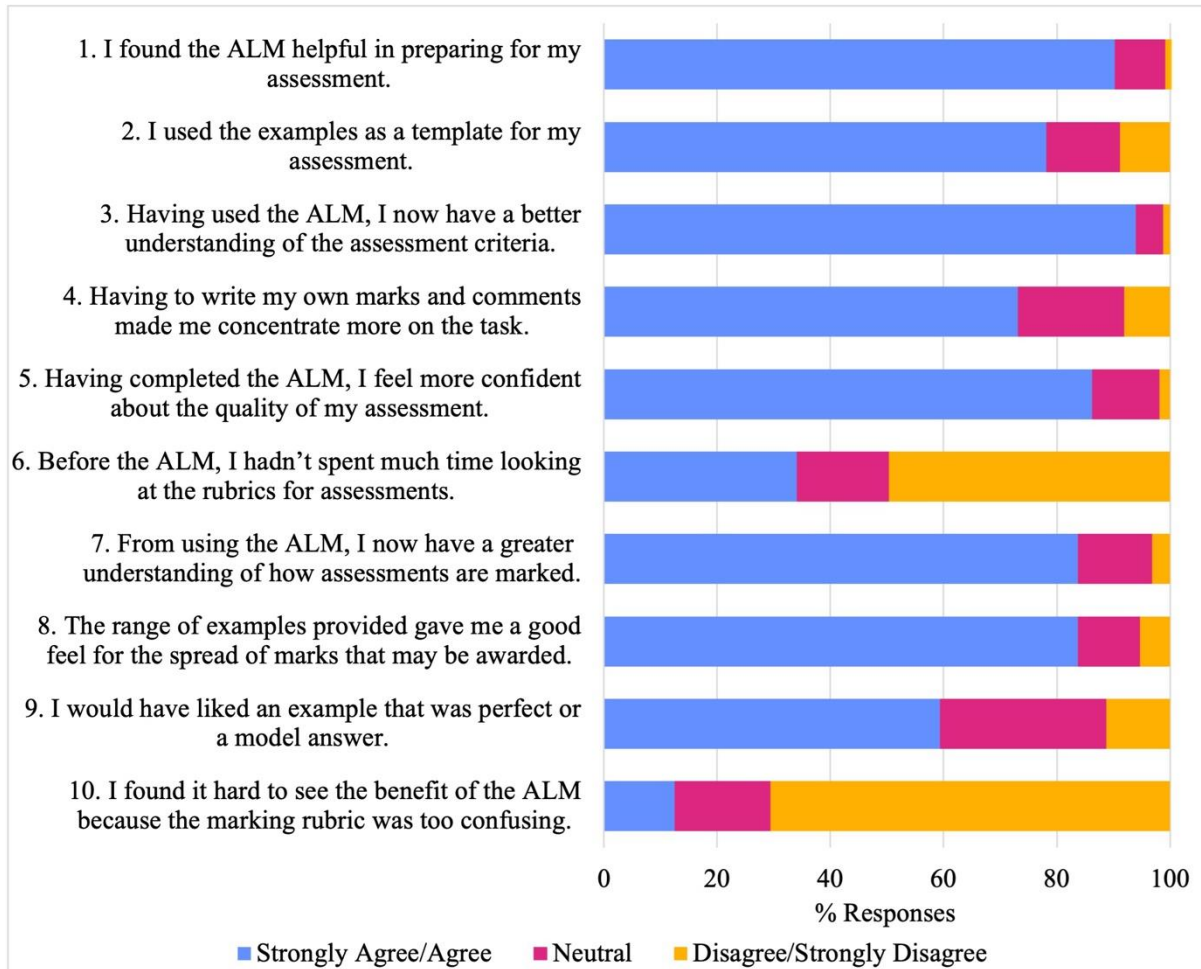


Figure 4. Students' perceptions (%) about their use of the ALM (N = 416).

After reading the definition of academic judgement as 'the skill of being able to self-evaluate your work, and the work of others', students were asked to indicate their agreement with four statements. Figure 5 illustrates that most students (92%) found using the ALM required academic judgement, the need to self-reflect, and be more considered in relation to the quality of their work when preparing their assessments. Half of the respondents (50%) suggested they gained further insight into past assessment performance after using the ALM. These findings were reinforced by student reflections: "[The ALM] gave me an opportunity to critically analyse academic work and insight into how our assessments are marked" and "[The ALM] forced me to perform the task that a marker would do...which in turn strengthened my ability to self assess my own work".

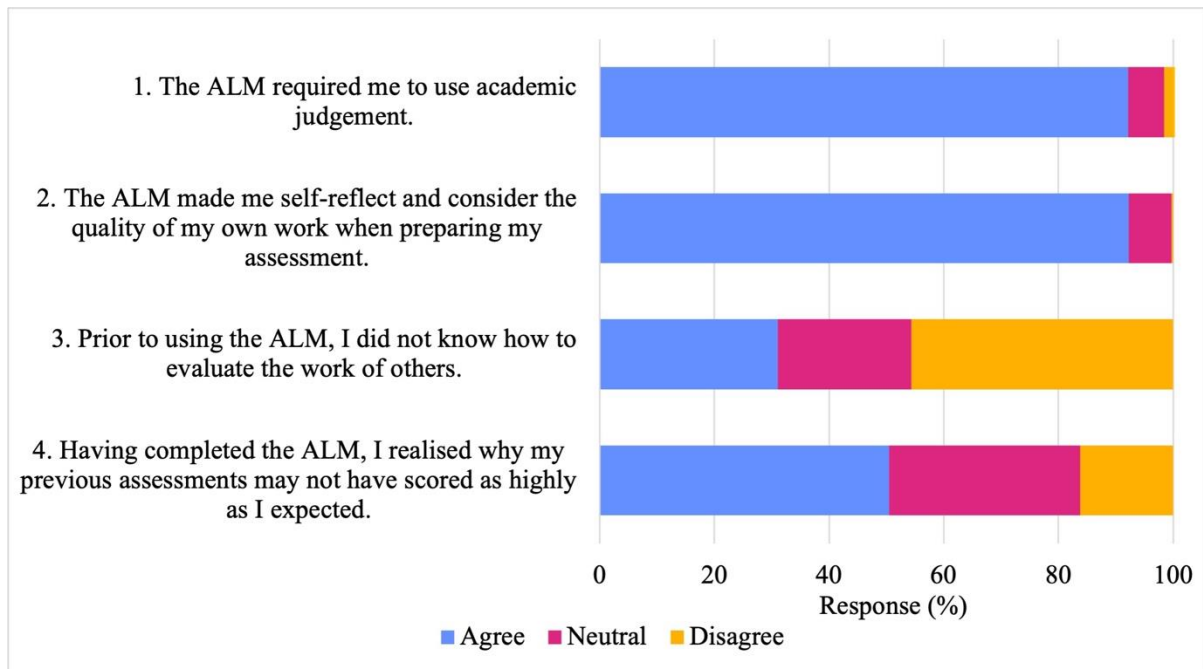


Figure 5. Students' perceptions of academic judgement and the ALM ($N = 309$; *Statement 1, $N = 318$).

Benefits to student assessment from completing the ALM (RQ4).

After examining student performance in ALM-associated assessments and comparing this against the broader student cohort, we observed that grades were mixed across different year levels, assessment types and subjects (Supplementary material), though in most subjects, the self-selected student survey participants achieved a 4-15% higher grade in their ALM-associated assessment with no discernible difference based on year level. Student's t-test comparing grades of survey participants against the general student cohort illustrated a statistically significant increase in assessment grade within most subjects (Supplementary material).

Student and staff reflections on future ALM use.

When asked about their future use of the module in relation to additional assessments in the subject, more than 80% of students agreed that they would spend more time reviewing the marking rubric as well as spending more time on self-reflection and self-assessment (Figure 6). Three quarters of students indicated that using the ALM changed the way they would approach their future assessments; with many students indicating they would like to have an ALM for other assessments in their subject (Figure 6).

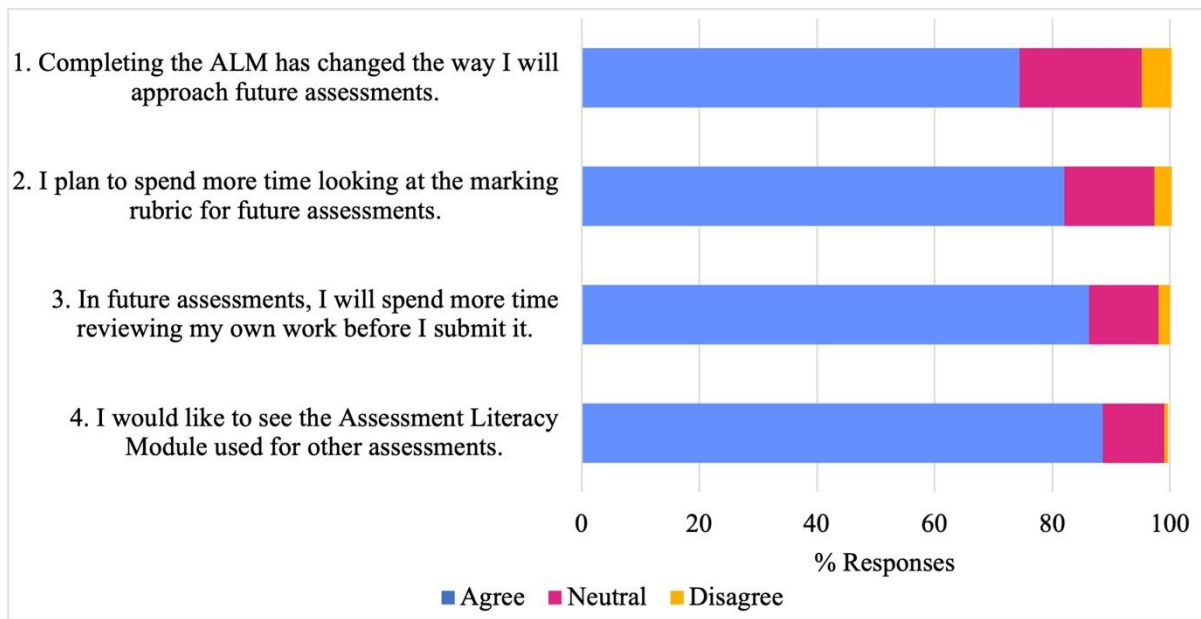


Figure 6. Students' perceptions of their future assessment behaviour ($N = 306$; *Statement 1, $N = 309$; *Statement 2, $N = 301$).

In response to the survey question 'Would you deploy the ALM in this subject next year?', all staff participants selected 'Yes' from the options given. However, 11 staff indicated that they planned to make changes in future iterations including: changes to teaching practice (for example, improvements in ALM instructions, promoting benefits to students better, increasing class time, updating and/or adding more examples periodically, and changing the weighting linked to ALM completion (for example, increase marks, make it compulsory/hurdle). More than half of staff participants (61%) indicated that they would deploy the ALM in other subjects in future semesters, while another 33% responded 'maybe'.

Students were also asked in the survey if there were changes that they would like to see in the ALM. Just under half ($n = 107$, 45%) indicated no changes were required. Table 2 summarises the most frequently suggested changes by students (only themes where three or more students indicated such a change are included).

Table 2. Most popular student themes for suggested changes to the ALM.

Theme	<i>n</i>	Illustrative quote(s)
1. Increase the number of examples and/or levels of examples, including higher-quality examples	55	<i>“I would like to see a greater spread of marks, I think it is good to show a passing example, a good/mid scoring example and then a high-scoring example.”</i> <i>“Seeing some lower marks would be good ... it would be nice to see what a pass (50%) would look like”</i>
2. Improve layout or navigation of ALM	17	<i>“Upgrading the layout will make it more desirable to do”</i>
3. Improve feedback and explanations of marks, criteria	12	<i>“Perhaps an extended feedback section breaking down each paragraph of the responses.”</i> <i>“For each criticism explain how much was penalised”</i>
4. Remove certain sections or include a “skip” option	8	<i>“Maybe an option to skip writing our own comments.”</i> <i>“The marking part is unnecessary, but the feedback should be kept.”</i> <i>“I feel like first impression marks...are not the most helpful”</i>
5. Use ALM in other subjects and/or in tutorials	5	<i>“I would like it to be used for more subjects and introduced earlier.”</i> <i>“More reference to it in tutorials – highlighting key points of the module”</i>
6. Give examples on how to provide comments and hints while marking	3	<i>“Having to make a comment is good however it can be difficult when you have not done this before ... maybe one report to review firstly that has assessor comments so you understand what to write in the comments”</i>

Discussion

This paper builds on data published in Willems-Jones et al., (2023) and addresses the limitations acknowledged in that paper regarding the low student response rate. In this extended study we report increased student survey participation as well as additional subjects and increased staff participation. The trends from the previous publication are upheld in this study extension, and both support our research hypothesis indicating that ‘Students who engage with the ALM perceive improvement in their ability to interpret assessment rubrics (contributing to assessment literacy) and in their ability to evaluate the quality and integrity (academic judgement) of assessment items, an outcome reinforced by staff perceptions’.

In response to RQ1, staff agreed that students who completed the ALM had a better understanding of their assessment expectations (Figure 2, Statement 2). Even though many staff indicated that setting up the ALM took a lot of time, staff reported that students who engaged with the ALM, evaluated examples, and compared their evaluation to the expert feedback comments better understood the assessment criteria (Figure 2, Statement 1). Staff also commented that the provision of feedback was easy to deliver to students and helpful for students to calibrate their academic judgement. Further positive aspects of the ALM identified by staff include increasing students’ agency and responsibility for their assessment. Overall, staff perceived the ALM to be a beneficial pedagogical tool, supported by the data indicating reasonable student engagement even in large cohort subjects.

A concern identified by staff in both the pilot (Willems-Jones et al., 2023) and this study was that examples in the ALM were likely to be plagiarised, particularly if these were high-scoring exemplars. The plagiarising of exemplars has also been reported by Bell, Mladenovic, & Price (2013). While 60% of students agreed that they would like more high-quality examples (Table 2, Theme 1) or an example that was a perfect/model answer (Figure 4, Statement 9), this does not necessarily equate to instances of academic misconduct. A high proportion of students (78%) indicated they used the examples as general templates for their assessment (Figure 4, Statement 2), a behaviour that has also been reported in the literature (Hendry & Anderson, 2013), and which is much broader in scope than direct plagiarism. Nevertheless, academics can minimise the risk of plagiarism by carefully selecting or constructing examples that have at least minor differences from the summative assessment task, have utilised different approaches and are of varying quality. It may be that the level of plagiarism observed during the use of the ALM was comparable to plagiarism levels for other non-ALM based assessments, as suggested by one of the staff respondents (S16), although this requires further investigation.

In addressing RQ2, we can confidently conclude that yes, staff do feel that ALM implementation improves the efficiency of their teaching practices. Figure 3, Statement 1 indicates that due to participation in the ALM, there were fewer student inquiries about the nature of the assessment. We also heard from staff that there was less work associated with communicating with students about assessment expectations. Furthermore, in addition to 56% of staff indicating there were improvements to academic workload related to assessment administration, many academics found the ALM useful in training casual markers.

The findings presented here illustrate that students felt the ALM had a positive influence on their understanding of assessment literacy, addressing the underlying research questions. We report here that this intervention also increases student confidence in their own judgement of quality. In terms of improving academic judgement (RQ3), most students agreed that using the ALM *required* them to use academic judgement and that completion of the ALM prompted them to self-reflect and ponder the quality of their work (Figure 5, Statements 1 and 2). Moreover, students reported that they had increased confidence about the quality of their work post-ALM completion (Figure 4, Statement 5). As demonstrated by Little, Dawson, Boud, and Tai (2024), evaluating improvements in student judgement is difficult and often achieved by measuring student confidence in their judgements. Similar to our findings, many researchers have observed that students report increased confidence in their own quality judgements following various interventions including using rubrics (Gyamfi, Hanna, & Khosravi, 2022), exemplars (Hawe & Dixon, 2017; To & Carless, 2016; To, Panadero, & Carless, 2022) and peer review (Ibarra-Sáiz, Rodríguez-Gómez, & Boud, 2020). The advantage of the ALM as an intervention, is that it incorporates all three approaches into one activity. Further work should include analysing the ALM metrics to compare student judgement of examples against the expert marking to see if this changes with each example attempted.

Reflecting on the second aspect of RQ3 (improvements in assessment literacy), our research indicates that student understanding of assessment literacy improved after engaging with their subject's ALM, and this is likely to lead to improvements in student learning (Smith et al., 2013). Our expanded research findings support our pilot study (Willems-Jones et al., 2023), and are consistent with literature such as that from Rust, Price, and O'Donovan (2003) indicating that student understanding of assessment criteria can be improved through interventions that explain marking criteria using exemplars. Our students report that they now have a better understanding of assessment criteria and how their work is graded (Figure 4, Statements 3 & 7), and students also agreed that the provision of a range of examples gave

them an appreciation for the distribution of marks (Figure 4, Statement 8), though not surprisingly, they wanted more exemplars. The survey responses indicate that by using the ALM, students also considered the quality of their own work, even without being explicitly asked to (Figure 5, Statements 1 and 2). This suggests some utility of the ALM lies in improving self-reflection and disposition towards assessment. The changes in student perceptions and attitudes following ALM use are all aspects of the multi-faceted concept of assessment literacy (defined by Hannigan et al. 2022), and therefore we can report that students think that using that ALM improved their assessment literacy.

In responding to RQ4, most students found that the ALM was helpful in preparing for their assessments, with many students using the examples as templates for the associated tasks (Figure 4, Statements 1 & 2). Students also reported that writing their own comments and marks within the ALM focussed their concentration on the task at hand (Figure 4, Statement 4). Moreover, student survey participants were observed to have higher grades (4-15 %) than the overall subject cohort in ALM-associated assessments in most subjects (Supplementary material), a finding consistent with literature which illustrates that student achievement can be improved through rubric use (Ragupathi & Lee, 2020; Reddy & Andrade, 2010). In addition, higher grades can be due to a student's increased awareness and their capacity for self-assessment and reflection when completing assessments (Boud et al., 2013). However, it is important to acknowledge a possible limitation that may partially confound this interpretation. As ALM use was self-selected and non-compulsory for most subjects, and student survey participants were also self-selected, the data presented may reflect a bias towards more proactive and engaged students, and thus improved grades cannot be definitively attributed to ALM use *per se*, but rather, they may simply reflect high-achieving and self-regulating students. In any case, student comments reveal that those who actively utilised the ALM were able to identify what makes a good assessment.

Not only did students perceive that the ALM was beneficial in preparing them for their associated assessments, using the module also changed their approach to future assessments (Figure 6, Statement 1), with most students intending to spend more time looking at marking rubrics and reviewing their own work (Figure 6, Statements 2 & 3). In line with self-regulated learning research (Panadero & Broadbent, 2018; Zimmerman & Moylan, 2009), the changes in how students approach assessment post-ALM suggest that students experienced improvements in self-regulated learning. It also suggests that by completing the ALM, many students' attitudes towards assessment have changed, further contributing to their improved assessment literacy; as described by Hannigan et al. (2022).

It is important to consider an additional potential limitation of the study, that not all students engaged with the opt-in ALM activity. Not only does this reduce the effective pool of survey respondents, but it also means that not all students received the benefits of the ALM, perceived or otherwise. Linking the ALM explicitly to assessment and delivered during class time may be an appropriate approach in the future, given that linked assessment appears to drive student engagement with the ALM (though not necessarily survey participation).

Conclusion

Our extended study shows positive outcomes in student perceptions of the ALM. The utility of the module has been demonstrated by its implementations in different disciplines, at different undergraduate levels and for different assessment tasks. The study indicates the ALM can be an effective tool to align staff and student conceptions of work quality and assessment

expectations across science and humanities disciplines. Furthermore, the ALM improves student assessment literacy by increasing student confidence in their academic judgement, and ability to interpret marking criteria, as well as improving understanding of contextualised assessment processes more generally. While additional research into the association of ALM engagement and further exploration into improvements in subject grades is required, the ALM is a useful pedagogical tool that provides unique feedback to aid in reaching consensus between staff and student expectations about assessments in university studies, including for large-scale subjects.

Acknowledgments

Development and implementation of the ALM was funded by UoM Learning & Teaching Initiative grants (multiple rounds). The Australian Council of the Deans of Science provided funding for the research assistant (CRM) and student participant remuneration. We would like to acknowledge Gordan Yau and the Learning Environments team at UoM for software development.

References

- Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research, and Evaluation*, 10(3). <https://doi:10.7275/g367-ye94>
- Bell, A., Mladenovic, R., & Price, M. (2013). Students' perceptions of the usefulness of marking guides, grade descriptors and annotated exemplars. *Assessment & Evaluation in Higher Education*, 38(7), 769-788. <https://doi:10.1080/02602938.2012.714738>
- Boud, D., Lawson, R., & Thompson, D. G. (2013). Does student engagement in self-assessment calibrate their judgement over time? *Assessment & Evaluation in Higher Education*, 38(8), 941-956. <https://doi:10.1080/02602938.2013.769198>
- 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:10.1080/02602938.2018.1522528>
- Gyamfi, G., Hanna, B. E., & Khosravi, H. (2022). The effects of rubrics on evaluative judgement: A randomised controlled experiment. *Assessment & Evaluation in Higher Education*, 47(1), 126-143. <https://doi:10.1080/02602938.2021.1887081>
- Hannigan, C., Alonzo, D., & Oo, C. Z. (2022). Student assessment literacy: indicators and domains from the literature. *Assessment in Education: Principles, Policy & Practice*, 29(4), 482-504. <https://doi:10.1080/0969594X.2022.2121911>
- Hawe, E., & Dixon, H. (2017). Assessment for learning: A catalyst for student self-regulation. *Assessment & Evaluation in Higher Education*, 42(8), 1181-1192. <https://doi:10.1080/02602938.2016.1236360>
- Hendry, G. D., & Anderson, J. (2013). Helping students understand the standards of work expected in an essay: Using exemplars in mathematics pre-service education classes. *Assessment & Evaluation in Higher Education*, 38(6), 754-768. <https://doi:10.1080/02602938.2012.703998>
- Hovardas, T., Tsvitanidou, O. E., & Zacharia, Z. C. (2014). Peer versus expert feedback: An investigation of the quality of peer feedback among secondary school students. *Computers & Education*, 71, 133-152. <https://doi:10.1016/j.compedu.2013.09.019>
- Ibarra-Sáiz, M. S., Rodríguez-Gómez, G., & Boud, D. (2020). Developing student competence through peer assessment: The role of feedback, self-regulation and evaluative judgement. *Higher Education*, 80(1), 137-156. <https://doi:10.1007/s10734-019-00469-2>
- Little, T., Dawson, P., Boud, D., & Tai, J. (2024). Can students' feedback literacy be improved? A scoping review of interventions. *Assessment & Evaluation in Higher Education*, 49(1), 39-52. <https://doi:10.1080/02602938.2023.2177613>
- Luo, J., & Chan, C. K. Y. (2023). Conceptualising evaluative judgement in the context of holistic competency development: results of a Delphi study. *Assessment & Evaluation in Higher Education*, 48(4), 513-528. <https://doi:10.1080/02602938.2022.2088690>
- Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. *Qualitative Health Research*, 26(13), 1753-1760. <https://doi:10.1177/1049732315617444>
- O'Donovan, B., Price, M., & Rust, C. (2008). Developing student understanding of assessment standards: A nested hierarchy of approaches. *Teaching in Higher Education*, 13(2), 205-217. <https://doi:10.1080/13562510801923344>

- Panadero, E., & Broadbent, J. (2018). Developing evaluative judgement: A self-regulated learning perspective. In D. Boud, R. Ajjawi, P. Dawson, & J. Tai (Eds.), *Developing Evaluative Judgement in Higher Education* (1st ed.). Routledge.
- Panadero, E., & Jonsson, A. (2020). A critical review of the arguments against the use of rubrics. *Educational Research Review*, 30, 100329. <https://doi:10.1016/j.edurev.2020.100329>
- Pintrich, P. R. (1995). Understanding self-regulated learning. *New Directions for Teaching and Learning*, 1995(63), 3-12. <https://doi:10.1002/tl.37219956304>
- Pintrich, P. R., & Zusho, A. (2002). The development of academic self-regulation: The role of cognitive and motivational factors. In *Development of achievement motivation*. (pp. 249-284). San Diego, CA, US: Academic Press.
- Ragupathi, K., & Lee, A. (2020). Beyond fairness and consistency in grading: The role of rubrics in Higher Education. In C. S. Sanger & N. W. Gleason (Eds.), *Diversity and Inclusion in Global Higher Education: Lessons from Across Asia* (pp. 73-95). Springer.
- Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, 35(4), 435-448. <https://doi:10.1080/02602930902862859>
- Rust, C., Price, M., & O'Donovan, B. (2003). Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment & Evaluation in Higher Education*, 28(2), 147-164. <https://doi:10.1080/02602930301671>
- Sadler, D. R. (2002). Ah!... so that's 'quality'. In P. Schwartz, Webb, G. (Ed.), *Assessment: Case Studies, Experience and Practice* (1st ed.). Routledge.
- Smith, C. D., Worsfold, K., Davies, L., Fisher, R., & McPhail, R. (2013). Assessment literacy and student learning: The case for explicitly developing students 'assessment literacy'. *Assessment & Evaluation in Higher Education*, 38(1), 44-60. <https://doi:10.1080/02602938.2011.598636>
- Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*, 76(3), 467-481. <https://doi:10.1007/s10734-017-0220-3>
- To, J., & Carless, D. (2016). Making productive use of exemplars: Peer discussion and teacher guidance for positive transfer of strategies. *Journal of Further and Higher Education*, 40(6), 746-764. <https://doi:10.1080/0309877X.2015.1014317>
- 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://doi:10.1080/02602938.2021.2011134>
- Willems-Jones, Munoz, C., Frankland, S., Saligari, M., Lim, S., Fox, J., McCarty, R., Fong, A., & Howe, P.D.L. (2023). Perceptions of an assessment literacy module to improve academic judgement – A pilot study. *Proceedings of the Australian Conference on Science and Mathematics Education, The University of Tasmania, 30 August – 1 September 2023*, pp. 83-90. <https://openjournals.library.sydney.edu.au/IISME/article/view/17419>
- Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where meta-cognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of Metacognition in Education* (pp. 299-315). Routledge.