

# Implementing the EAT Framework to Enhance Assessment Literacy and Feedback Practices

Active Learning in Higher Education

1–20

© The Author(s) 2025



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/14697874251360577

[journals.sagepub.com/home/alh](https://journals.sagepub.com/home/alh)

Melody Pattison<sup>1</sup> , Sara M. Pons-Sanz<sup>1</sup>  
and Mia Waters<sup>1</sup>

## Abstract

This study investigates the effects of the implementation of the EAT (Equity, Agency, Transparency) framework on the assessment and feedback practices of a second-year undergraduate module running at a Russell Group university in the UK, and how the concepts underpinning the framework help to improve students' assessment literacy and self-criticality. To achieve this, in the academic years 2022 and 2023 students were surveyed on their understanding of the marking criteria (in the categories of *Content and Coverage*, *Presentation and Style*, *Organisation and Structure*, *Originality and Critical Thinking*, and *Literature and Referencing*) and their confidence in evaluating and critiquing their own work (*Critique Own Work* and *Implement Feedback*) at the beginning and end of the semester. As part of the intervention, students were required to complete a summative three-step peer review process, where they had to provide feedback on a peer's submission and critically reflect on their own work. Finally, students were asked for qualitative comments on what they found helpful about the process, and were required to identify specific areas for feedback in their final assessment.

Following the intervention, the survey results showed that students' understanding and confidence increased in all categories, with statistical significance in most categories between the first and second surveys in both years. Qualitative comments also indicated how students felt they had become more self-critical and could apply what they had learnt to future assessments, while students who identified specific areas for feedback demonstrated ownership of their learning. Thus, the study recommends the explicit teaching of the assessment criteria with practical opportunities to put this knowledge into practice in order to increase students' overall assessment literacy and confidence.

## Keywords

assessment, feedback, EAT framework, self-criticality, marking criteria

<sup>1</sup>School of English, Communication and Philosophy, Cardiff University, Cardiff, UK

## Corresponding author:

Melody Pattison, School of English, Communication and Philosophy, Cardiff University, John Percival Building, Colum Drive, Cardiff, CF10 3EU, UK.

Email: [pattisonm@cardiff.ac.uk](mailto:pattisonm@cardiff.ac.uk)

## Introduction

Within a UK Higher Education context, students have shown difficulty with understanding assessment marking criteria, particularly in a written form without explicit discussion (Graham et al., 2022; O'Donovan et al., 2001; Rust et al., 2003). The resulting impact from this struggle tends to be lower attainment in academic achievement and, in turn, decreasing self-confidence.

The aim of this study is to investigate the effects of the implementation of the EAT (Equity, Agency Transparency) framework (Evans, 2016) on the assessment and feedback practices of a second-year undergraduate module at a UK University, and how the concepts underpinning the framework help to improve students' assessment literacy and self-criticality. Case studies on the implementation of the framework have been reported on since its inception (e.g. Balloo et al., 2018; Evans et al., 2019; Zhu et al., 2018), and the current study situates its findings within the existing literature on EAT, as well as other research in the areas of assessment literacy and feedback in Higher Education.

In section 'Assessment of this article, we outline relevant studies in the field of Higher Education assessment practices focused on interventions in the areas of assessment, marking criteria and feedback. In section 'The EAT Framework and Our Adaptation', we give an overview of the EAT framework and our adaptation, before we detail our methodology in analysing the effectiveness of our intervention (section 'Methodology') and present our analysis (section 'Analysis of Results') and conclusions (section 'Conclusions and Recommendations').

The research questions this study wished to answer are:

- 1) What is the impact of the explicit teaching of assessment criteria and the summative three-step peer review on students' assessment literacy?
- 2) What is the impact of the explicit teaching of assessment criteria and the summative three-step peer review on students' feedback practices and self-criticality?

## Assessment

Formative assessment as a component of Assessment for Learning (AfL) is associated with strategies which contribute to self-regulatory behaviours and meta-cognitive self-monitoring (Hawe & Dixon, 2017), which in turn enhance learning outcomes by giving students more agency over their own learning. It is highly relevant to how EAT (Evans, 2016) aims to aid students' overall assessment literacy. Implementing formative feedback elements into summative assessment for larger cohorts may also help students to be more engaged, self-critical and motivated to improve in their future work (Broadbent et al., 2018). Overall, the findings from the studies discussed through this section highlight the dialogic use and evaluation of exemplars, peer and self-assessment, rubrics, audio explanations and feedback as formative ways to improve students' metacognitive thinking.

### *Peer Assessment and Self-Assessment*

Both peer assessment and self-assessment can engage students in acts of critical evaluation and judgement, developing their capacities as self-regulated learners (Nicol et al., 2014); however, Orsmond et al. (2002) found that peer assessment produced more objective judgement than self-assessment where marking criteria co-produced by staff and students, exemplars and formative feedback were implemented. Nicol et al. (2014) show how students revisit, rethink and update their work as a result of engaging in peer review activities, and that students feel giving feedback is slightly more valuable than receiving it due to the activation of critical thinking skills involved in giving feedback. Subsequently, students develop the ability to think more critically, both of others'

work and their own. Additionally, Theising et al. (2014) report that receiving peer feedback has a positive impact on student self-confidence. Therefore, it is evident that both parties benefit from such a process, particularly the providers of the feedback.

Scott et al. (2014) outline the 'shadow module' approach to student learning. This approach requires the involvement of three different types of 'peers' who run collaborative, peer-mediated learning activities. The 'peers' can be grouped into 'true peers' (students on the module in which the intervention is run), 'near-peer' (students from a previous cohort) or 'far peers' (post-graduate students). Typically, 'true peers' may be experiencing the same difficulties, 'near peers' know what is challenging having previously completed the module and 'far peers' might have wider subject knowledge and teaching experience (Amici-Dargan et al., 2020). It is suggested that sessions are most productive with little input from academic staff and led by either 'near peers' or 'far peers'. Students tended to disengage when there was too much input from academic staff, and they found challenging material more accessible when explained by a peer (Scott et al., 2014). Amici-Dargan et al. (2020) propose that 'shadow modules' encourage learning through their discursive and collaborative nature with peers rather than academic staff, and benefit the three possible types of participants: students who are either highly engaged, less engaged or 'lurkers' (those who do not actively participate but do view the resources). However, despite the benefits of the sessions, some students do not participate at all, and typically the number of active participants is around 10% to 20% (Scott et al., 2014). In our intervention, a student who previously went through this assessment process on our module offered peer support to students in the most recent cohort; however, in line with Scott et al.'s (2014) findings that active participation in this process is low, no students in the cohort engaged with her.

Stommel (2024) argues for self-evaluation and for students to frequently reflect upon their own learning, and Cowan and Harte (2023) suggest that a programme should 'concentrate attention on needs for development identified by the students' (p.51). An online self-assessment tool may also be beneficial to students' final performance in assessment tasks. Ibabe and Jauregizar (2010) found a positive correlation between the use of a voluntary online self-assessment tool and stronger academic performance. Making self-assessment summative may encourage more students to complete it, thereby enhancing its effects. Nicol et al. (2014) recommend the use of peer review followed by self-review using the same criteria in order to promote self-criticality.

### *Marking Criteria and Rubrics*

Marking criteria in Higher Education in the form of rubrics can be useful when closely aligned with the assignment (Jonsson, 2014) and used to promote self-regulatory behaviours (Andrade & Du, 2005); however, the importance should be on explaining individual marking criteria as opposed to criteria as a whole (Orsmond et al., 1997, 2000). This practice may involve introducing the entire rubric, but focusing on the individual criteria and what students must do in order to achieve certain scores in each category. For example, the criteria for the grades associated with each individual category must be clearly worded and explained, and students must have an understanding about what the categories refer to. Price and Rust (1999) explored academic staff's perspective on a common assessment criteria grid; their analysis was followed by O'Donovan et al. (2001), who considered the student perspective on the grid. The consensus from the staff perspective was that, although the grid was useful in providing some clarity on grade descriptors, it failed to improve student outcomes; the language used in the grid was too complicated, and it would be more beneficial if used in conjunction with a discussion of the grid components before assessment submission (Price & Rust, 1999). The student perspective aligned with this conclusion; students who chose not to use the grid stated their reasons were due to it having a lack of detail and the need for academic

staff to explain it (O'Donovan et al., 2001). Students overall felt that the criteria were vague and that, although the grid had potential, it was of limited practical use unless presented within a multi-faceted approach including explanations, exemplars, and opportunities for discussion (O'Donovan et al., 2001). Similarly, Graham et al. (2022) suggest that assessment-specific marking criteria (as opposed to generic criteria), engagement sessions and electronic provision of feedback can make assessment and feedback more effective, while Fraile et al. (2017) and Orsmond et al. (2000) advocate for the co-creation of criteria and rubrics, as co-creation promotes explicit discussions on the assessment criteria (see also Boyd & Cowan, 1985). Without such discussions, it is possible that some students may have a different understanding of the criteria compared to their tutors and peers (Orsmond et al., 1997), and it is important that this misunderstanding is avoided or corrected.

Rust et al. (2003) devised an intervention in which undergraduate students attended a workshop which included small group discussions, and tutor comparison of exemplars and explanation of criteria; students also marked exemplars themselves in preparation for the workshop. Students who attended this workshop performed significantly better on their assignment than students who did not. In addition, students were required to submit a self-assessment form with their assignment, and it was found that participants in the workshop were more likely to underestimate their grade than non-participants (who were more likely to overestimate); this is indicative of the participants' ability to think more critically than non-participants. These results align with those of Orsmond et al. (1997), who also found that students who did well were more likely to be self-critical and under-mark their work when completing self-assessment.

In addition to rubrics, exemplars are effective at enhancing students' understanding of marking criteria (Orsmond et al., 2002). Recommendations for how exemplars may be introduced include pre-annotated (or accompanied by a video explanation) exemplars (Broadbent et al., 2018), engagement sessions (Graham et al., 2022), online exemplars (e.g. through a virtual learning environment) giving students easy and frequent access (Handley & Williams, 2011), and the discussion of exemplars with students (Carless & Chan, 2017; To & Carless, 2016; To & Liu, 2018). In particular, online exemplars increase self-regulatory behaviour, and dialogic use helps clarify concepts. A challenge with dialogic use is the power-relations evident in the teacher-student interactions; however, with reduced teacher authority, discussion of exemplars can promote improvements in students' agentic participation (To & Liu, 2018). If encouragement to use peer dialogue (Carless & Chan, 2017) is implemented, or students present their own exemplar analysis before any teacher-led discussion (To & Liu, 2018), dialogic use of exemplars can be very effective in enhancing students' self-regulation and academic judgement skills.

It is necessary to acknowledge that researchers such as Stommel (2024), Spurlock (2023), and Clark and Talbert (2023) advocate for alternative forms of assessment, such as ungrading, which minimises the pressure and importance of final grades. As with methods such as co-creation, this approach was unsuitable for our module (as the criteria are standard across our degree programme); however, our study places a high level of importance on the active role of feedback and self-evaluation in much the same way as the above-named researchers.

## Feedback

At the centre of feedback issues are the questions of what students think the purpose of feedback is, and what makes effective feedback. Dawson et al. (2019) addressed these questions in a qualitative analysis of the difference between student and staff perspectives. Their results showed that 90% of students in their study comprising 400 students from two Australian universities thought the purpose of feedback is to help improvement (89% of 323 staff responded the same). Both educators and students were predominantly focused on improvements to work and ability to produce

work, rather than an improvement in the ability of students to evaluate work; the fact that an improvement in the ability to self-evaluate is not claimed by either educators or students is also a limitation. In terms of what makes for effective feedback, there was more divergence between staff and students. While 53% of staff believed it was the overall design of the feedback process which promoted effectiveness, only 17% of students mentioned this. In contrast, students instead thought the contents of the comments themselves impacted feedback effectiveness (84% of students compared to 34% of staff). These findings show the need for staff and students to be aligned on both the purpose of feedback and how to gauge its effectiveness. As only a relatively small majority of staff believed the design of the feedback process, rather than the comments themselves, promoted effectiveness, the question is raised that staff as well as students may need to think more carefully about the nature and purpose of feedback.

Although Orsmond et al. (2005) found student perception of feedback to be centred around relevant themes such as its role in enhancing motivation and learning, encouraging reflection and clarifying understanding, the results from Dawson et al.'s (2019) study show a need for greater clarity about how feedback is used and understood in practice.

Winstone et al. (2017) identified some barriers preventing student use of feedback. These included confusion and frustration around the language used in feedback, or not understanding how to use feedback to improve future performance, thereby showing difficulty in translating feedback into action. In order to overcome these barriers, student active engagement needs to be considered. Strategies such as feed-forward, audio feedback and dialogic feedback have been shown to be effective in eliciting this type of engagement (Graham et al., 2022; Hill & West, 2019; Tam, 2021).

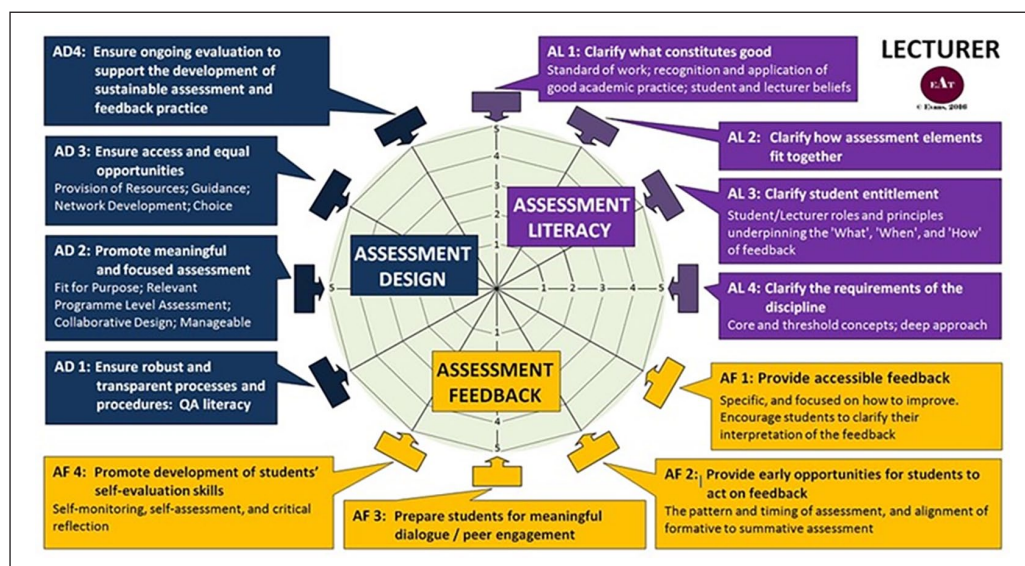
Feed-forward refers to feedback which is directed at how a student can improve in future assessments, lessening the issue of the lack of transferability of feedback, which in turn can lead to student disengagement. Hill and West (2019) found that after students on an undergraduate ecology course undertook a feed-forward process, the mean course mark improved from pre to post-intervention by 7%. The intervention involved a multi-step process of students discussing their draft essay face-to-face with the course tutor, discussing and grading exemplars with their peers, writing self-reflections of their essay progress, and finally self-assessing their own essay. Students self-reported that they evaluate their learning better after completing the process. Preceding Hill and West, Taras (2001, 2003) also advocates for integrated peer and/or tutor feedback within any process of student self-assessment.

Regarding the use of exemplars (see Section 'Feedback'), Tam (2021) found that students had mostly positive perceptions of exemplar-based dialogic feedback, but that some students still held the perception that feedback should be transmitted by teachers in a one-way delivery method. There is therefore evidently the need to help students reconceptualise their meaning of feedback and change their mindset from one of transmission of learning to reciprocal learning. This type of dialogic feedback could work best when combined with interventions such as the feed-forward meetings investigated by Hill and West (2019), in which students are encouraged to reflect on their own work while in a face-to-face meeting with their course tutor.

## **The EAT Framework and Our Adaptation**

The EAT framework (Evans, 2016) is a research-informed approach focused on improving assessment practices to develop students' skills in assessment literacy and self-regulation. It contains three core dimensions of practice (Assessment Literacy, Assessment Design and Assessment Feedback), which are further divided into an additional 12 subdimensions which can aid educators in improving students' overall assessment literacy (see Figure 1). Our approach aimed to involve as many of these components as possible.





**Figure 1.** The EAT framework (Evans, 2016, p.16).

There have been few previous case studies on the implementation of the EAT framework in Higher Education (Balloo et al., 2018; Evans et al., 2019; Zhu et al., 2018). Zhu et al.'s (2018) case study focuses on lecturers' engagement with assessment and feedback practices as a way of enhancing student assessment literacy, and considers the experiences of academics in implementing assessment initiatives to promote student self-regulation based on the EAT framework. Overall, they concluded that, to advance students' engagement with assessment and feedback, support must be invested in lecturers' pedagogic research literacy in the aforementioned areas.

Balloon et al. (2018) introduced an assessment brief template built on the principles of the EAT framework across first-year undergraduate modules at the University of Surrey. Their aim was to explore patterns between students' perceptions of the effectiveness of the assessment brief templates in developing their assessment literacy and self-reported self-regulation. Through the use of textual analysis software, the researchers identified words belonging to specific linguistic domains in response to open-ended questions designed to investigate how students perceived the intervention. The findings suggested that negative language was associated with lower levels of self-regulation, and positive language associated with higher levels.

Finally, Evans et al.'s (2019) project focused on reducing differential learning outcomes, in particular for Black and Asian minority ethnic (BAME) students, and students from lower socio-economic backgrounds, through training staff in three Higher Education Institutions (Southampton, Surrey and Kingston) to implement EAT principles. Their case studies across these institutions found increased co-production of assessment and levels of self-regulation, with the ultimate recommendation that a more nuanced understanding is needed of how individual difference and variables combine to impact students' learning outcomes. While co-production can be a valuable tool for improving assessment literacy, it is important to note that as part of our study it was not possible to co-create the marking criteria with the students, because the criteria are standard in our school.

Our study focuses on a co-taught second-year undergraduate module currently running at a Russell Group university in the UK. It focuses on stylistics and covers a number of theoretical

frameworks for the linguistic investigation of a wide range of texts, both literary and non-literary. The number of students enrolled on the module usually ranges from 45 to 60, and its contact hours each week consist of a 2-hr lecture and a 1-hr seminar.

Following student feedback on the need to better understand the assessments and the assessment criteria, changes to the assessment in the module began to be implemented in the 2021/22 academic year. The module initially consisted of three assessment tasks: a peer feedback task, a textual analysis and an essay. The textual analysis and essay were initially worth 50% each, until the participation in the peer review task became a summative exercise in 2021/2022. From this year onwards, the percentage each assessment task was worth was adjusted: the peer review task at 10%, the textual analysis at 50% and the essay at 40%.

Although the types of assessment tasks were not changed, the approach to them was updated following the components of the EAT Framework, beginning in the 2021/2022 academic year:

- 1) Assessment Feedback (Evans, 2016), comprising AF2, AF3 and AF4 (see Figure 1): The peer feedback task was initially an optional, formative task. It was originally developed so as to give students a chance to practise their textual analysis skills, as this was a type of assessment that they were not used to, particularly as the module brings together students from different degrees. The challenge with the peer feedback task was that many students either chose not to engage with it, or questioned why other students provided feedback for them instead of the module leaders (similar comments to Tam's, 2021 findings).

To address these issues, and as advocated by researchers such as Brown et al. (1997), the peer-review process was turned into a three-step summative task similar to Nicol et al.'s (2014) approach. Students were required to engage with all three components to gain full marks: (1) submit a short textual analysis on a text provided to them; (2) provide feedback on a classmate's submission on the basis of the assessment criteria that would be used to mark their work (see 'Assessment Literacy', below); (3) evaluate their own short analysis on the basis of the same criteria. The process was summative only in that students were required to complete it in order to receive a mark on a 0%/100% basis. They did not receive reviews from the lecturers; the point of the exercise was to develop their own self-criticality and not to receive feedback from us. Therefore, despite students' participation in the process being required as part of the module, in practice it functioned as a more facilitative, formative exercise. To take inclusivity and confidentiality into account, all feedback was anonymous: students did not know who reviewed their work, or whose work they were reviewing. Anonymous student comments following the process indicated that it was helpful for the task to be made summative, as they were then forced to engage with it and consequently could see the benefits.

Students were also asked to include a reflective paragraph at the end of the textual analysis outlining how/whether they found the peer review process helpful (and, specifically, 'how participating in the peer review process has helped [them] to write this textual analysis'), and a paragraph at the end of the essay where they had to identify specific aspects of their work (on the basis of the marking criteria) they wanted feedback on and the reasons for their choices. These paragraphs were intended to help the students further reflect on their work and to identify areas themselves in which they would like more support. In line with Cowan's (2020) emphasis on reflective practices linking with graduate attributes, it was important that the students understood how self-criticality is relevant not only to their modules and assessments, but also transferable to other areas and skills they can carry with them throughout their lives.

- 2) Assessment Literacy (Evans, 2016), comprising AL1 and AL2 (see Figure 1): Prior to completing the peer review task (for the first 4 weeks), part of the seminar each week was devoted to a specific marking criterion, so students were familiar with them before undertaking the first summative task (the textual analysis). The descriptors for each mark band in connection with the marking criteria have always been a part of the module guide, but not necessarily explicitly discussed. This part of the process follows the results from previous research (O'Donovan et al., 2001; Price & Rust, 1999; Rust et al., 2003), which found that explicit discussion of the assessment criteria contained in marking grids was necessary for students to fully understand and familiarise themselves with the criteria (see Section 'Marking Criteria and Rubrics'). Markle (1964, 1978) and Tiemann and Markle (1990) also advocate for the use of examples and non-examples; given the time limitations, it was not possible to include these, but students were provided with sample answers from previous years and encouraged to talk to the lecturers if they had any questions.
- 3) Assessment Design (Evans, 2016), comprising AD1, AD2 and AD3 (see Figure 1): For the textual analysis, students initially were asked to choose from a series of texts, to which they could apply their chosen theoretical framework. In the updated assessment, the set texts were replaced with the opportunity for students to choose any text they wish, with them being especially encouraged to choose a text they had studied in another module, so that they could then see the connections between different theoretical approaches applied to the same text.

The aim of this intervention was for the assessment process to be more meaningful to the students and directly linked to their own learning, with the ultimate goals of improving students' assessment literacy, feedback, and self-criticality. The EAT framework formed the basis of changes to how assessment practices worked in the module. Table 1 provides an overview of the pre and post-EAT framework interventions.

It is, of course, important to acknowledge and emphasise the avoidance of over-assessment, which is a concern widely commented on in Higher Education (Bennett, 2000; Tomas & Jessop, 2019; Yu et al., 2023). However, these interventions are intended to enhance the overall quality of assessments, not to add unnecessary burdens on students and markers. Engaging with the tasks should ultimately lead to a greater understanding of assessment practices, thus translating into higher student achievement. These assessments aim to provide a balance between not over-assessing while at the same time giving students opportunities to practise and not making their assessment dependant on a single piece at the end of the module.

**Table 1.** Overview of the Changes Made to the Assessment in Our Module and the Alignment to the EAT Framework.

Pre-EAT	Post-EAT	Aspects covered
General seminar content	Seminars consist of module content alongside explicit discussion of marking criteria	AL1, AL2
Optional peer review	Summative three-step peer review process	AF2, AF3
Textual analysis on set texts	Textual analysis on text of students' choice plus reflective paragraph	AD1, AD2, AD3, AF4
Essay	Essay plus identification of criteria for specific feedback	AF4



# Methodology

This case study follows a mixed-methods approach; it combines quantitative analysis of survey data and assessment results, with a qualitative analysis of student comments. In order to judge the impact and address the research questions, the various types of data considered were:

- 1) The average mark for the two main assessment tasks (textual analysis and essay) across four different academic years: 2020/2021 (before implementation of EAT framework), 2021/2022 (first year of implementation of EAT framework) and 2022/2023 (second year of implementation of EAT framework).
- 2) Pre- and post-intervention survey results in the academic years 2022/2023 and 2023/2024.
- 3) Comments included in reflective paragraphs by students in the assessments for the 2022/2023 and 2023/2024 academic years.

In order to see the impact of the intervention, from 2022 during face-to-face seminars, students were surveyed on their understanding of the marking criteria and their confidence in evaluating and critiquing their own work. Students were asked at the beginning of the module to rate their understanding of the expectations of the marking criteria (*‘Content and Coverage’*, *‘Presentation and Style’*, *‘Originality and Critical Thinking’*, *‘Use of Literature and Referencing’*), and to rate their ability to critique their own work and to implement the feedback they receive in future pieces of work. The ratings were coded on a five-step scale from Very Poor to Very Good. At the end of the module, once students had had an opportunity to learn further about and reflect on the marking criteria as part of the weekly intervention and the submission of the three pieces of summative assessment, they were asked to complete the same survey again. It should be noted that the numbers of students responding to the survey differed between the beginning and end of the semester in each year (see Table 2). The discrepancy between the number of respondents can be attributed to the second survey being conducted at the end of the semester, by which time student attendance had dropped due to a number of reasons, for example, going home for holidays or working on assignment deadlines in other modules. Worsening attendance levels towards the end of the semester are not unusual (cf. Rodgers, 2001; van Blerkom, 2001; van Walbeek, 2004).

The data were then statistically analysed using the Wilcox test in R to form the basis of a quantitative discussion. This test was chosen because the results being compared consisted of two ordinal variables, and the aim was to find out whether the typical value at the end of the intervention was higher than at the beginning. In addition, the Wilcox test did not require adjusting for having different numbers of responses at the beginning and end of each semester. Qualitative results, in the form of student comments on the overall usefulness of the process, are also considered in Sections ‘Methodology’ and ‘Analysis of Results’

Finally, the overall assessment results from the 2020/2021 academic year (the year before the implementation of the EAT framework) and the following academic years (when the EAT framework began to be implemented) were compared.

**Table 2.** Number of Survey Respondents.

Year	Survey 1 responses (n)	Survey 2 responses (n)
2022	53 (87% of cohort)	28 (46% of cohort)
2023	45 (88% of cohort)	32 (63% of cohort)

## Analysis of Results

This section considers the quantitative survey results and qualitative comments on the effect of the implementation of EAT principles in the module. In addition, the overall marks for the module's assessments from the 2020/2021 academic year through to the 2023/2024 academic year are compared quantitatively in order to judge whether the implementation of the intervention helped to improve academic attainment.

### Quantitative Results from Survey Data

The Wilcoxon test was applied to the results of the survey questions as outlined in Section 'The EAT Framework and Our Adaptation' (see Table 3). The results show that students' reported understanding and confidence increased in all categories following the intervention, and there is statistical significance in most categories between the first and second surveys in both academic years 2022/2023 (hereafter '2022') and 2023/2024 (hereafter '2023').

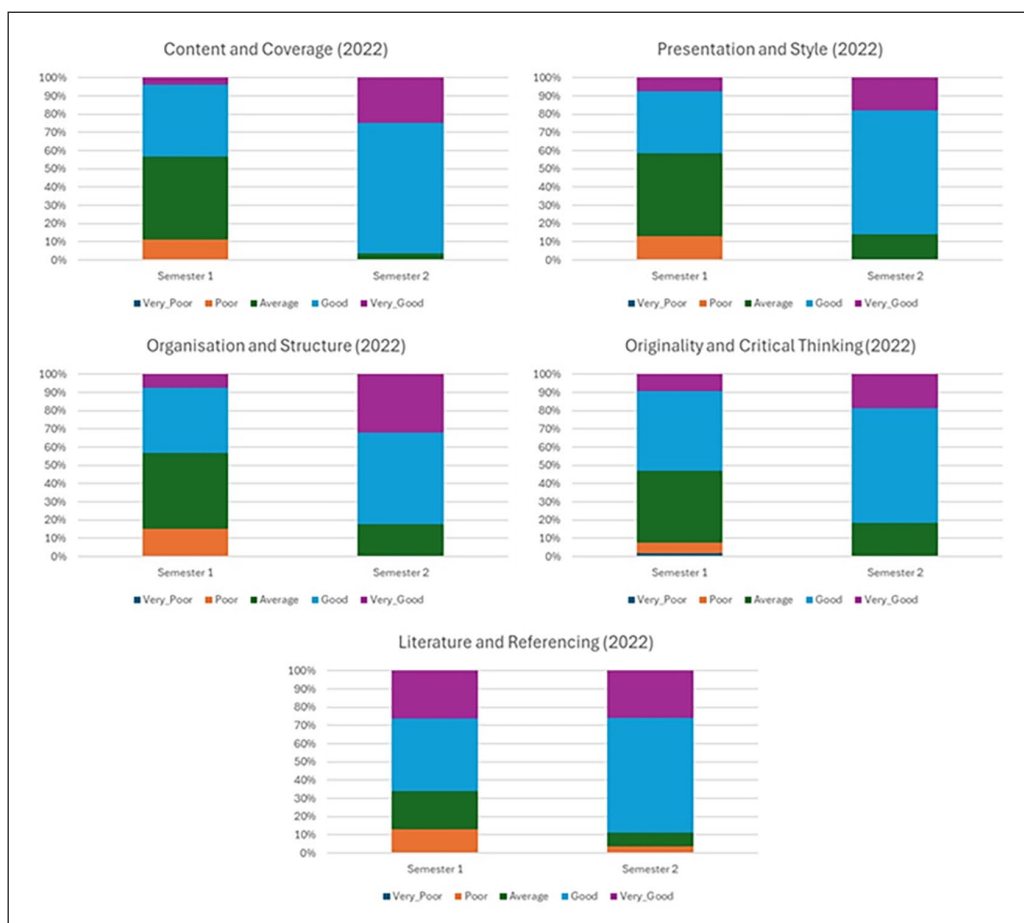
With reference to the marking criteria, the categories showing the highest statistical significance in increased understanding in both 2022 and 2023 are 'Content and Coverage', 'Presentation and Style', and 'Organisation and Structure', where in all cases  $p < .01$ . 'Originality and Critical Thinking' in 2022 also shows statistical significance where  $p < .05$ . For the additional categories relating to students' reported ability to critique their own work and implement feedback received, the results for 'Critique Own Work' and 'Implement Feedback' are statistically significant in 2022 with  $p < .01$ , and 'Critique Own Work' in 2023 is also significant with  $p < .05$ . The results for 'Implement Feedback' were not found to be statistically significant in 2023; however, the direction of change still showed an increase in students' perceived ability (see Figure 5). The only category whose results did not show statistical significance in either 2022 or 2023 is 'Literature and Referencing'; however, again, the direction of change is still increasing, and students generally reported a good understanding of this category in the initial survey.

Figures 2 to 5 provide a visual representation of the increase in understanding of each category in 2022 and 2023.

Figures 2 to 5 show that in both years there is a clear shift in students' reported understanding of the marking criteria and confidence in their ability to critique their own work and implement feedback. Two categories, 'Literature and Referencing' and 'Critique Own Work' in 2022 did show a decline in students rating their knowledge as 'Very Good'; however, for 'Critique Own Work', no students rated their knowledge as 'Poor' or 'Very Poor' by the end of the semester, and the Wilcoxon test confirms a statistically significant increase in the direction of change.

**Table 3.** Statistical Analysis Using the Wilcoxon Test.

Category	Wilcoxon $p$ value (2022)	Wilcoxon $p$ value (2023)
Content and coverage	1.1E-06	8.22E-06
Presentation and style	0.000193	0.000962
Organisation and structure	0.000125	0.000894
Originality and critical thinking	0.010783	0.069879
Literature and referencing	0.194139	0.160501
Critique own work	0.001014	0.024431
Implement feedback	0.002153	0.546869



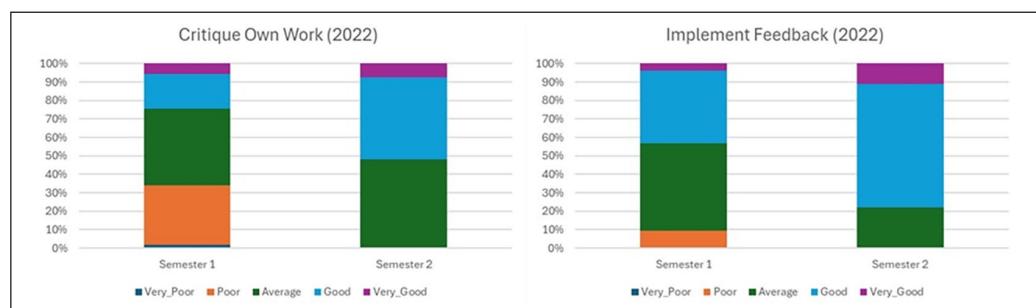
**Figure 2.** Graphs showing increase in understanding of marking criteria (2022).

In both years, the largest increase in reported understanding was in the marking criteria category ‘Content and Coverage’, with the majority of students rating their knowledge as ‘Good’ or ‘Very Good’ by the end of the semester. ‘Presentation and Style’ and ‘Organisation and Structure’ also show similar ratings in both 2022 and 2023. As the ‘Implement Feedback’ category showed statistical significance in 2022, but not 2023, it could be an area of increased focus in future iterations of the module.

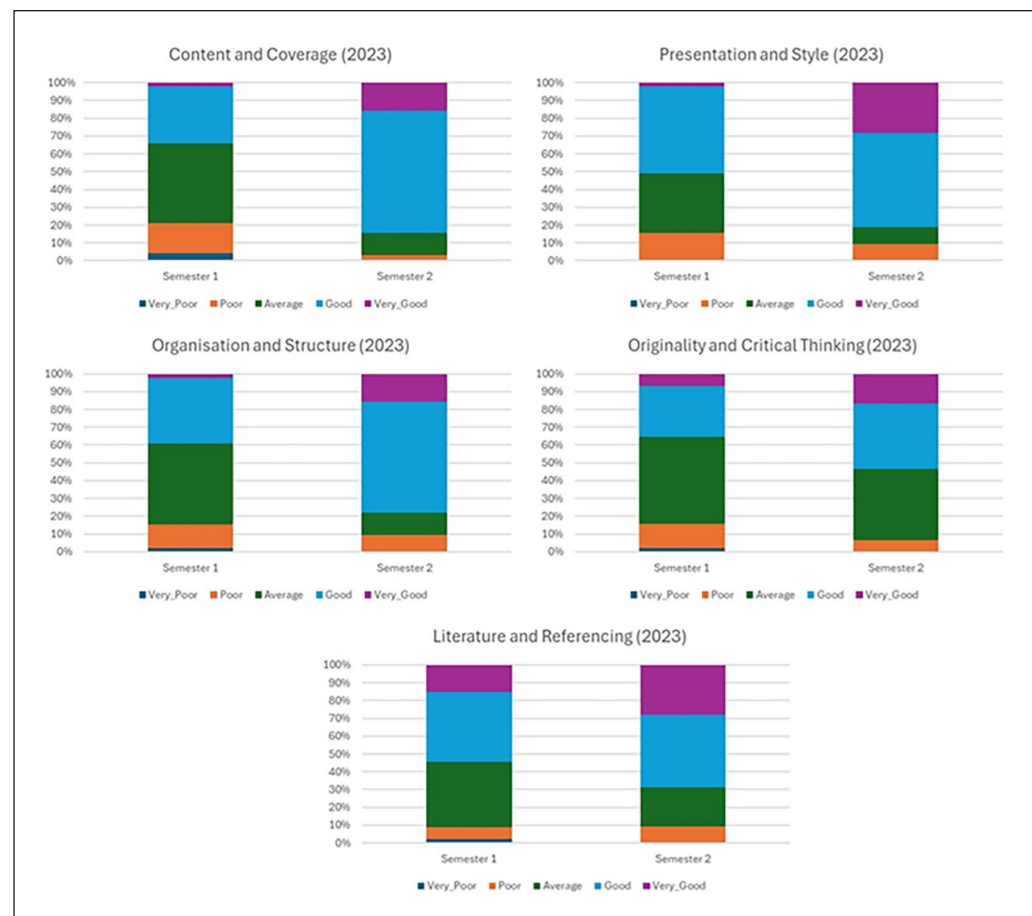
Overall, the quantitative analysis of results lends support to the effectiveness of the interventions, with statistically significant findings in the majority of categories, and an increase in the direction of change in all of them.

### *Qualitative Results from Student Comments*

When submitting their main textual analysis, students were required to include a paragraph reflecting on their participation in the peer review process, commenting on the overall usefulness of the latter and the explicit discussions on the marking criteria for their work on the textual analysis. Comments from both 2022/2023 and 2023/2024 cohorts were overwhelmingly positive and indicated several themes:

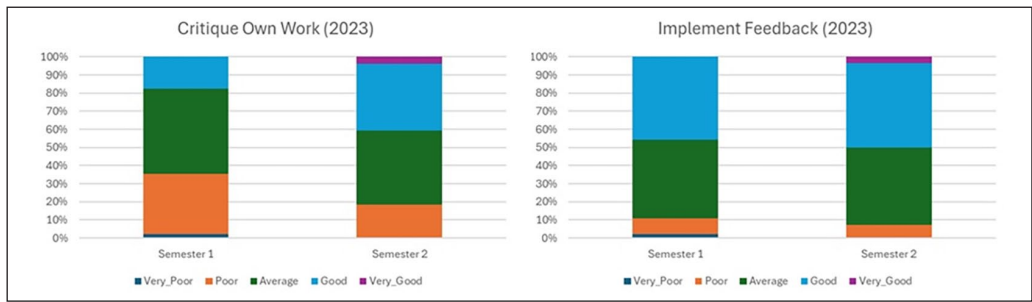


**Figure 3.** Graphs showing increase in confidence in abilities (2022).



**Figure 4.** Graphs showing increase in understanding of marking criteria (2023).

- 1) Students reported that they had learnt to be more self-critical and consider other viewpoints:



**Figure 5.** Graphs showing increase in confidence in abilities (2023).

Overall, it was a helpful exercise, as it made me more critical of my own writing and reading someone else's, allowing me to look at my work with a more critical perspective.

Reviewing my own work after the peer review was helpful because I could look at my work with consideration of a viewpoint which I was not entirely aware of before.

2) Students reported that they had a greater understanding of how their work was marked:

The peer review process really helped me understand the key criteria I need to be hitting for top marks. Things that I spotted missing in my peer's work made me reconsider whether I had missed the same criteria in my own writing. On the other hand, my peer excelled at hitting other criteria which made me see how I could then implement that into my own work.

By participating in the peer review task, it has helped me to focus on each marking criteria (*sic.*) more so than I would have. I have been more aware especially of my presentation and style by adopting a formal voice. . . Taking part in the peer review task has enabled me to understand how I am marked and what I need to work on to improve the overall presentation of my work.

3) Students reported that they thought about how they can implement what they learned for future assessments:

I can see how self-evaluation can be really beneficial and I will definitely use it to improve my academic writing in the future.

By analysing work and perhaps seeing a potential for improvement places that point at the forefront of my mind when I am constructing an essay in the future.

4) Students reported that they found it useful receiving feedback from a peer:

Through the use of the peer-review process, I found it easier to improve my writing as it was read by someone who was of a similar skillset to mine and . . . had to follow the same guidelines as me.

The most helpful part . . . was the feedback. It was good to get someone else's opinion as I made some mistakes in my analysis.

It is interesting, however, to note this final comment. Although the student has reported finding the opinion of a peer to be helpful, they also seem to have missed a key intention of the



process. Comments such as these are suggestive that some students are still not aware that the key target is their own ability to critique their work, even though the key purpose of the exercise was made explicit a number of times because of similar comments in previous years (e.g. that students wanted feedback from lecturers). This could be taken as a reminder that one cannot assume that students fully engage with all the sessions in a module (lectures and seminars), and key messages need to be discussed throughout. However, other student comments indicated they critically engaged with what their peers had excelled at, and could implement these aspects in their own work (see themes 1 and 2). Through this process, students could thus identify their strengths and weaknesses in a way similar to a very effective constructive feedback approach facilitated by Cowan and Harte (2023), which evidenced the emergence of highly metacognitive self-evaluation skills.

As part of the third assessment task, the essay, students were required to identify specific aspects of their work they would like feedback on, although they were also given a marking grid with an indication of how well they had performed in each of the marking criteria. This component was intended to give students more ownership over their own learning and reflect critically on potential areas for improvement. Most students identified specific aspects of the assessment criteria, and some referred to the feedback they received as part of the previous textual analysis, looking to further improve on aspects that had been commented on as part of that piece of assessment:

After reflecting upon the feedback I have received from my textual analysis, I would find it very much useful to gain deeper feedback on the areas: originality and critical thinking, content and coverage, and my use of literature and referencing. I feel these are the areas that I try and engage with consistently and appropriately throughout my work; however, I do feel as though I do not fully engage with all aspects of these areas which therefore limits the quality produced in my essay. After reading previous feedback, it is apparent to me that although I exemplify an interesting analysis of the subject discussed in my essay, I do not fully explore the language or the secondary sources used in my work which I think limit my credibility.

The above answer shows a good example of a student who has engaged fully with the process and is aware of the areas on which they need to focus. They justify why they think these are the areas on which they'd appreciate further feedback, demonstrating their self-criticality. Other responses also showed similar awareness of weaker points:

I would like to receive feedback on how well my essay reads because I have a tendency to ramble and include irrelevant information in my writing. I would also like feedback on my use of terminology, because I find terminology quite confusing. Finally, I would like feedback on my use of references, particularly whether I've embedded them into the text correctly because I worry about citing references incorrectly or it being unclear which reference I'm using.

I struggle with formatting my ideas in a concise way that avoids waffling until I eventually get to the point I'm trying to make.

While doing the essay I found that I was unsure of how to approach the engaging with secondary sources aspect. I used ideas from secondary reading to support my analysis, but I didn't really incorporate them entirely into an analysis I felt was my 'own voice', and so would like feedback on how I could have used secondary reading in a more effective way.

Comments such as these suggest that students have developed a sense of self-awareness and self-criticality, and they are able to identify areas of their work on which they would appreciate further

guidance. As mentioned above, this exercise was intended to give them ownership over their learning and to make the feedback they receive work for them in a meaningful way. Although these comments show that some students are dependent on further assistance from the lecturers when it comes to the application of certain aspects of the criteria, they have progressed to the stage where they have an awareness of where their weaknesses lie. Further self-reflective practices and a willingness to act on feedback should ultimately lead them to independently assess how they have understood the criteria.

### *Average Marks for Assessment Tasks*

Table 4 shows the breakdown of the average mark (where a mark of 40% is a pass and the higher marks do not tend to be higher than 78%) for both the textual analysis and the essay, from the 2020/2021 academic year (pre-intervention) and the following three academic years (post-intervention). The data do not include resits or students who received a mark of 0% by not submitting the assessments. The results show that, although student self-reported confidence in their own ability and understanding has clearly increased, it is yet to be reflected in the results of the assessment tasks following the peer review process. It should also be noted that the lecturers on the module have remained the same across the years, with a similar split in the marking of the scripts.

However, we can see some impact from the intervention when considering the overall breakdown of marks (see Tables 5 and 6). For the essay, we can see a slight increase in the number of students receiving marks in the First range (>70%) in the second year of the intervention compared to previous years. However, this could also be attributed to the fact that the lecturers have been explicit about what the focus needs to be for the essay, and how it differs from the textual analysis. In previous years (before this intervention), some students submitted work that was closer to a textual analysis than an essay. As a result of this, the lecturers have been careful since to detail the specific differences between what is required for a textual analysis and what is required for an essay.

**Table 4.** Average Assessment Results for the Textual Analysis and Essay Across Four Academic Years.

Year	Textual analysis average mark	Essay average mark
2020/2021	64.65	58.77
2021/2022	60.55	60.26
2022/2023	60.72	60.43
2023/2024	61.72	59.42

**Table 5.** Breakdown of Marks for the Textual Analysis.

Year	<40 (%)	40–49 (%)	50–59 (%)	60–69 (%)	70–79 (%)	80–89 (%)	>90 (%)
2020/2021	0	2.1	16.3	61.2	20.4	0	0
2021/2022	7.5	3.8	18.9	52.8	17	0	0
2022/2023	4.9	11.5	19.7	44	19.7	0	0
2023/2024	2	2	30	52	14	0	0

**Table 6.** Breakdown of Marks for the Essay.

Year	<40 (%)	40–49 (%)	50–59 (%)	60–69 (%)	70–79 (%)	80–89 (%)	>90 (%)
2020/2021	0	16.7	33.3	39.6	10.4	0	0
2021/2022	0	14	20	54	12	0	0
2022/2023	0	4.9	37.7	39.3	18	0	0
2023/2024	2	8	42	38	10	0	0

## Conclusions and Recommendations

The most significant finding of this study is the effect that the implementation of EAT practices (Evans, 2016) has on students' self-reported understanding of how they are marked, confidence in their ability to implement feedback, and self-criticality. In the 2 years the surveys were run, the direction of change in students' self-reported understanding and confidence in their abilities was to increase, with statistically significant results being reported in at least 1 year for all categories except 'Use of Literature and Referencing'. The increase in understanding of 'Content and Coverage', 'Presentation and Style', and 'Organisation and Structure' were the most significant in both years the survey was run ( $p < .001$  for both years). Qualitative comments submitted as part of the textual analysis also indicated how students felt they could apply what they had learnt to future assessments. Comments submitted as part of the essay submission were intended to give students ownership of their feedback by asking the lecturers what they would like specific feedback on, and many students demonstrated a clear and thoughtful understanding of their own strengths and weaknesses, showing the ability to be self-critical.

It is clear that, although the results of the intervention do not seem to yet be reflected in the marks for the textual analysis, they are starting to become apparent with the essay. One explanation for this may be that, although some students seemed to miss the focus of the essay, treating it like another textual analysis, others are able to put everything that they have learned regarding the marking criteria and the feedback they received in the textual analysis into practice in the final piece of assessment. In future years, we are looking to replace the essay with another textual analysis, because what we want is for the students to be able to apply the frameworks, and the distinction between the textual analysis and the essay seems to be causing unnecessary problems. The hope is that, and when taking into account student comments about application to future study, this self-criticality and reflectiveness should then continue to be developed throughout the remainder of the students' degree programme (and beyond), consequently resulting not just in continually improving assessment results, but transferable skills and highly sought-after graduate attributes (Cowan, 2020). The focus on self-criticality has been fully integrated into the revalidated degree programmes this module is part of, to be taught from September 2025 onwards. Therefore, while this element is specific to this module at this point in time, it should be better integrated into the School's degree programmes in the future.

The study does of course need to consider a number of limitations. Firstly, each year produced a different cohort of students, so they cannot always be directly comparable as they each bring with them their own set of skills, strengths and weaknesses. Secondly, the pre- and post-intervention surveys were conducted during face-to-face seminars. This approach resulted in unequal numbers of respondents, as by the end of the semester fewer students were attending classes. It also cannot be guaranteed that the same students answered both surveys; however, there was overlap because most students in the seminars attended all of them. Finally, we should also acknowledge that we have approached learning as being a cognitive effort (Sayalı et al., 2023); thus, student motivations (and subsequent achievement) may differ depending on the individual.

Overall, in spite of these limitations, this study has provided us with answers to the two research questions in Section 'Introduction'. With regard to Research Question 1, the results suggest that implementing features of the EAT Framework (Evans, 2016) has a significant impact on students' claimed understanding of marking criteria. This impact has been measured through the students' survey responses at the beginning and end of the semester (where the Wilcoxon test showed statistically significant evidence of increased reported understanding in most categories), alongside slightly increased assessment results each year we implemented the intervention, in contrast with those years when we did not. With regard to Research Question 2, the results suggest an increase in students' reported confidence in their abilities at implementing feedback and being self-critical, again evidenced through statistically significant survey results using the Wilcoxon test and qualitative comments provided by students in the reflective paragraph. The recommendation is that explicit discussions on assessment criteria can be utilised to increase students' understanding of what they are actually expected to produce (see also O'Donovan et al., 2001; Price & Rust, 1999; Rust et al., 2003). Additionally, a summative peer and self-review task on a pass/fail basis helps students to apply the marking criteria and engage with self-criticality in preparation for the main pieces of assessment (cf. Nicol et al., 2014). These recommendations are supported by the significance of the survey results, but it would also be useful to follow students' academic achievements over a longer period of time in order to make a judgement on how successful the intervention is with further opportunities to put its intentions into practice.

## Acknowledgements

The authors would like to thank Sean Roberts for his invaluable help with statistical testing, and Cardiff University for its financial support as part of its summer internship programme.

## ORCID iD

Melody Pattison  <https://orcid.org/0009-0009-8499-0497>

## Ethical Considerations

The research project was reviewed and given a favourable opinion by the ENCAP School Research Ethics Committee at Cardiff University on 26th June 2023.

## Consent to Participate/Publish

Written informed consent, including consent to publish, was obtained by all participants.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This work was supported by the Learning and Teaching Academy at Cardiff University through their summer internship programme.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

## Data Availability Statement

The data that support the findings of this study are available from the authors but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available.

## References

- Amici-Dargan, S. L., Moorcroft, A. M., & Rutherford, S. M. (2020). Shadow modules and student-led online learning communities: Supporting learning, engagement, and student empowerment within, and across, institutions in higher education. In Dennis, C., Abbott, S., Matheson, R., & Tangney, S. (Eds.), *Flexibility and pedagogy in higher education* (pp. 79–96). Brill.
- Andrade, H. L., & Du, Y. (2005). Student perspective on rubric-referenced assessment. *Educational & Counseling Psychology Faculty Scholarship*, 10(3), 1–11.
- Balloo, K., Winstone, N., & Norman, M. (2018). *Evaluation of a large-scale inclusive assessment intervention: A novel approach to quantifying perceptions about assessment literacy*. University of Surrey.
- Bennett, M. (2000). Assessment to promote learning. *The Law Teacher*, 34(2), 167–174.
- Boyd, H., & Cowan, J. (1985). A case for self-assessment based on recent studies of student learning. *Assessment and Evaluation in Higher Education*, 10(3), 225–235.
- Broadbent, J., Panadero, E., & Boud, D. (2018). Implementing summative assessment with a formative flavour. *Assessment & Evaluation in Higher Education*, 43(2), 307–322.
- Brown, G., Bull, J., & Pendlebury, M. (1997). *Assessing student learning in higher education*. Routledge.
- Carless, D., & Chan, K. K. H. (2017). Managing dialogic use of exemplars. *Assessment & Evaluation in Higher Education*, 42(6), 930–941.
- Clark, D., & Talbert, R. (2023). *Grading for growth: A guide to alternative grading practices that promote authentic learning and student engagement in higher education* (1st ed.). Routledge.
- Cowan, J. (2020). ‘Students’ evidenced claims for development of abilities arising from linked reflection-on-action and reflection-for-action. *Reflective Practice*, 21(2), 159–170.
- Cowan, J., & Harte, C. (2023). Towards a partnership model for developing core abilities. *International Journal for Students as Partners*, 7(1), 39–54.
- Dawson, P., Henderson, M., Mahoney, P., Phillips, M., Ryan, T., Boud, D., & Molloy, E. (2019). What makes for effective feedback: Staff and student perspectives. *Assessment & Evaluation in Higher Education*, 44(1), 25–36.
- Evans, C. (2016). *Enhancing assessment feedback practice in higher education: The EAT framework*. University of Southampton.
- Evans, C., Zhu, X., Balloo, K., Hughes, A., & Bright, C. (2019). *Maximising student success through the development of self-regulation*. Office for Students.
- Fraile, J., Panadero, E., & Pardo, R. (2017). Co-creating rubrics: The effects on self-regulated learning, self-efficacy and performance of establishing assessment criteria with students. *Studies in Educational Evaluation*, 53, 69–76.
- Graham, A. I., Harner, C., & Marsham, S. (2022). Can assessment-specific marking criteria and electronic comment libraries increase student engagement with assessment and feedback? *Assessment and Evaluation in Higher Education*, 47(7), 1071–1086.
- Handley, K., & Williams, L. (2011). From copying to learning: Using exemplars to engage students with assessment criteria and feedback. *Assessment & Evaluation in Higher Education*, 36(1), 95–108.
- Hawe, E., & Dixon, H. (2017). Assessment for learning: A catalyst for student self-regulation. *Assessment & Evaluation in Higher Education*, 42(8), 1181–1192.
- Hill, J., & West, H. (2019). Improving the student learning experience through dialogic feed-forward assessment. *Assessment & Evaluation in Higher Education*, 45(1), 82–97.
- Ibabe, I., & Jauregizar, J. (2010). Online self-assessment with feedback and metacognitive knowledge. *Higher Education*, 59, 243–258.
- Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment & Evaluation in Higher Education*, 39(7), 840–852.
- Markle, S. M. (1964). *Good frames and bad: A grammar of frame writing*. John Wiley & Sons, Inc.
- Markle, S. M. (1978). *Designs for instructional designers*. Stipes Publishing Company.
- Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education*, 39(1), 102–122.
- O’Donovan, B., Price, M., & Rust, C. (2001). The student experience of criterion-referenced assessment through the use of a common criteria assessment grid. *Innovations in Learning and Teaching International*, 38(1), 74–85.



- Orsmond, P., Merry, S., & Reiling, K. (1997). A study in self-assessment: Tutor and students' perceptions of performance criteria. *Assessment & Evaluation in Higher Education*, 22(4), 357–368.
- Orsmond, P., Merry, S., & Reiling, K. (2000). The use of student derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education*, 25(1), 23–38.
- Orsmond, P., Merry, S., & Reiling, K. (2002). The use of exemplars and formative feedback when using student derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education*, 27(4), 309–323.
- Orsmond, P., Merry, S., & Reiling, K. (2005). Biology students' utilization of tutors' formative feedback: A qualitative interview study. *Assessment & Evaluation in Higher Education*, 30(4), 369–386.
- Price, M., & Rust, C. (1999). The experience of introducing a common criteria assessment grid across an academic department. *Quality in Higher Education*, 5(2), 133–144.
- Rodgers, J. R. (2001). A panel-data study of the effect of student attendance on university performance. *Australian Journal of Education*, 45(3), 284–295.
- 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.
- Sayali, C., Heling, E., & Cools, R. (2023). Learning progress mediates the link between cognitive effort and task engagement. *Cognition*, 236, Article 105418.
- Scott, J. L., Moxham, B. J., & Rutherford, S. M. (2014). Building an open academic environment: A new approach to empowering students in their learning of anatomy through 'Shadow Modules'. *Journal of Anatomy*, 224, 286–295.
- Spurlock, S. (2023, March 15–18). *Improving student motivation by ungrading* [Conference session]. 54th ACM Technical Symposium on Computer Science Education V. 1, Toronto, ON, Canada (pp. 631–637).
- Stommel, J. (2024). Ungrading: An introduction. *Pedagogy*, 24(3), 327–340.
- Tam, A. C. F. (2021). Undergraduate students' perceptions of and responses to exemplar-based dialogic feedback. *Assessment & Evaluation in Higher Education*, 46(2), 269–285.
- Taras, M. (2001). The use of tutor feedback and student self-assessment in summative assessment tasks: Towards transparency for students and for tutors. *Assessment & Evaluation in Higher Education*, 26(6), 606–614.
- Taras, M. (2003). To feedback or not to feedback in student self-assessment. *Assessment & Evaluation in Higher Education*, 28(5), 549–565.
- Theising, K., Wu, K., & Sheehan, A. H. (2014). Impact of peer assessment on student pharmacists' behaviors and self-confidence. *Currents in Pharmacy Teaching and Learning*, 6(1), 10–14.
- Tiemann, P. W., & Markle, S. M. (1990). *Analyzing instructional content: A guide to instruction and evaluation*. Morningside Press.
- 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.
- To, J., & Liu, Y. (2018). Using peer and teacher-student exemplar dialogues to unpack assessment standards: Challenges and possibilities. *Assessment & Evaluation in Higher Education*, 43(3), 449–460.
- Tomas, C., & Jessop, T. (2019). Struggling and juggling: A comparison of student assessment loads across research and teaching-intensive universities. *Assessment & Evaluation in Higher Education*, 44(1), 1–10.
- van Blerkom, M. (2001). Class attendance in undergraduate courses. *The Journal of Psychology*, 126(5), 487–494.
- van Walbeek, C. (2004). Does lecture attendance matter? Some observations from a first year economics course at the University of Cape Town. *South African Journal of Economics*, 72(4), 861–882.
- Winstone, N. E., Nash, R. A., Rowntree, J., & Parker, M. (2017). 'It'd be useful, but I wouldn't use it': Barriers to university students' feedback seeking and recipience. *Studies in Higher Education*, 42(11), 2026–2041.
- Yu, D., Swanepoel, C., Sims, D., & Anciano, F. (2023). Do assessment loads affect student academic success? An investigation at an economics department at the undergraduate level. *South African Journal of Higher Education*, 37(5), 237–259.
- Zhu, X., Ford, N., Spencer, V., Lotti, E., & Evans, C. (2018). *Building pedagogical research literacy in assessment practices within Higher Education*. University of Southampton.

**Author biographies**

**Melody Pattison** is a Lecturer in Linguistics at the School of English, Communication and Philosophy, Cardiff University, where she teaches courses on sociolinguistics, phonetics, and stylistics. She has also published on vowel variation and change within a regional Dutch dialect.

**Sara M. Pons-Sanz** is a Professor of Historical Linguistics at the School of English, Communication and Philosophy, Cardiff University. Besides pedagogy, she is interested in the impact of medieval multilingualism on the development of English, historical stylistics and glossography.

**Mia Waters** recently graduated from Cardiff University with a degree in English Literature and Language and currently holds a graduate role. Her academic interests include self-driven learning and self-regulatory best practices, which she now applies in her corporate role.