

Enhancing the assessment literacy of distance learning students: A case study

Kate Whittington *
*Bristol Medical School,
University of Bristol, Bristol, UK.*
k.whittington@bristol.ac.uk,

Ellayne Fowler
*Centre for Health Sciences Education,
University of Bristol, Bristol, UK.*
ellayne.fowler@bristol.ac.uk

Oscar Cordero Llana
*Bristol Medical School,
University of Bristol, Bristol, UK.*
oscar.corderollana@bristol.ac.uk,

Abstract

Actively engaging students with mark schemes has been shown to improve their subsequent assessment performance. However, such activities rely on a degree of tacit knowledge transfer which can be difficult for large cohorts or distance learning students. This case study considers whether online discussion around a generic task can help distance learners improve their assessment literacy. Distance learning students engaged in a task to design their own mark scheme and use it to mark a series of 'mock papers' designed to mimic a range of common marking situations e.g. fail, including irrelevant material, plagiarism. The marks of the first essay these students subsequently undertook did not significantly differ irrespective of whether they had participated in this task. However, the range of marks obtained by those that did not engage in the task was significantly more variable than by those who had engaged (F-variance test, $p=0.0003$), suggesting the task may help to bring consistency to the performance and expectations of students from a range of academic background. The views of students on this task were positive with three quarters agreeing that it had

¹ * Corresponding Author

improved their understanding of mark schemes and over 60% (63%) confirming it would help them to grade their own work in the future. This case study illustrates that interactive campus-based tasks designed to improve students' assessment literacy can be converted into e-learning sessions that enable the explicit exploration of tacit knowledge.

Keywords: assessment criteria, reflection, distance-learning, assessment literacy, mark scheme.

Introduction

A student's ability to improve their assessment mark has been suggested to be aided by increasing their 'assessment literacy' (Rust et al., 2003). This was described by Smith et al., (2013) as a series of actions ranging from understanding the purpose of the assessment, identifying good elements of their work or areas for improvement through to understanding of the marking process and, hence marking schemes. Criterion-referenced assessments are widely used across the Higher Education sector and are suggested to aid assessment reliability (Price & Rust, 1999) and the maintenance of academic standards (Lucas and Webster, 1998). Despite the limitations of this approach, associated with individual interpretation of criterion and grade boundaries (Webster et al., 2000), it is valued by students for providing guidance on what is expected within an answer (O'Donovan et al., 2001). However, most criterion-referenced assessment schemes are complex and only the most motivated students are likely to engage with them in the absence of discussion and support (Price & Rust, 1999 and O'Donovan et al., 2001).

Learning has been suggested to involve the need for the transfer of tacit as well as explicit knowledge (Anderson & Ostman, 2015, Asher & Popper, 2019). Tacit knowledge is based more around experience and its transfer involves activities such as imitation, practice and discussion. Rust et al., (2003) suggested that increasing student's understanding of criterion-referenced marking and hence their assessment

literacy, by activity and discussion-based tasks, significantly improved student's assessment performance. Interestingly this intervention appeared to have long-term benefits with improved assessment performance still being present a year later (Rust et al., 2003).

A range of methodologies to facilitate the transfer of tacit knowledge associated with understanding assessment criteria can be incorporated into the curricula of campus-based programmes (O'Donovan et al., 2008). However, with increasing numbers of students in higher education (Higher Education Statistics Agency, 2019) such teaching methods are resource intensive and not suitable for all modes of study. Distance learning in particular poses challenges. However, identifying ways in which distance learners can improve their understanding of assessment criteria would not only improve their personal academic performance but could also be useful for learners in more traditional, campus-based settings, thereby easing the resource burden.

Distance learning is often associated with specific benefits (Salmon, 2005) but one specific challenge is the transfer of tacit knowledge (Ubon & Kimble, 2002). The use of online interactive technologies, such as discussion forums or social networking, allows group work and discussion to occur and therefore makes the transfer of tacit knowledge an achievable goal (Panahi et al., 2012) for remote learners. The aim of this project was to assess the suitability of a series of online tasks and discussions in helping a cohort of distance learning students to better understand the purpose of criterion-referenced marking schemes and how they could use them to improve their own assessment performance.

Methods

Summary

The University of Bristol has offered a distance learning Master of Science (MSc) programme in Stem Cells and Regeneration (<http://www.bristol.ac.uk/study/postgraduate/2020/health-sciences/msc-stem-cells-regeneration/>) since 2009. Graduates from a range of academic backgrounds (biomedical sciences, medicine, dentistry, medicine, veterinary sciences, etc.) partake in the programme from different locations worldwide. To enhance the assessment literacy of students on the programme an intervention was designed that required newly enrolled students to review the programme's marking scheme and use it to develop their own mark scheme for a non-academic task of taking a photograph of a famous Bristol landmark. An online discussion forum was provided to allow students and staff to discuss their schemes which they then used to assess some pre-generated pieces of work. Following this exercise, participants were invited to share their views through a task evaluation survey. The mark schemes generated by the students, were analysed linguistically in terms of their use of adjectives and compared to that of the programme's scheme. Assessment marks from this cohort of students were reviewed in order to determine the long-term impact on assessment performance.

Running of the Assessment Literacy Task

Students (n=18) who enrolled in the 2016/17 academic year were invited to participate in a voluntary task to help improve understanding of the programme's criterion-referenced marking scheme. All students were distance learners from a variety of geographic locations. All learning, teaching, discussion and support occurred via remote methods, including online lectures, videos and explanations of practical techniques, web-based discussion forums, email, skype and telephone conversations.

A specific area on the programme's virtual learning environment (Blackboard) was set up explaining the task and its purpose. As a first step, students were asked to review the programme's criterion-referenced marking scheme before being asked to develop their own mark scheme that could be used for the following assignment: '*Submit an original photograph showing the beauty of the Clifton Suspension Bridge, making sure you capture the full architectural detail of the whole bridge in landscape view*'.

We purposefully selected a topic that was not linked to the academic content of this programme so that students would not feel they were being judged on their subject knowledge. The topic, the Clifton Suspension Bridge, is one of Bristol's most famous tourist attractions and significant information is publicly available online thus enabling students who had never visited Bristol to participate and feel more engaged with the City. Students were specifically asked to consider the difference in criterion that would result in a fail (< 50%), pass (50% - <70%) merit (60 - <70%) and distinction ($\geq 70\%$) level grade.

Students were asked to upload their own marking criterion for the specified task to an online discussion board to which all students, irrespective of whether they participated with the task or not, were enrolled. Communication over the discussion board was asynchronous, which was expected as the majority of learners on this programme are part-time students. Therefore a few days were given for individuals to complete their mark scheme and upload it. However, a tutor responded promptly to all student mark schemes as they were uploaded with questions and specific discussion points were used to encourage the whole group to reflect on their own and each other's schemes. For example, they were asked what they found difficult and why they thought that was. In order to help students consider the different requirements for Bachelors (level 6) and Masters (level 7), comments and links to specific information study were included, such as the UK's Quality Assurance Agency's (QAA) Framework for Higher Qualification Standards (<https://www.qaa.ac.uk/docs/qaa/quality-code/qualifications-frameworks.pdf>).

Once students had uploaded their individual mark scheme and there had been discussion on this aspect of the task, they were subsequently asked to use their mark scheme to assess a series of photographs that had been uploaded to Blackboard.

These were examples of work that might be submitted for the assignment, ie photos of Clifton Suspension Bridge. Photographs were used to cover the range of possible marking bands, such as fail, pass, merit and distinction. In addition, photos were taken to show specific marking issues, such as plagiarism, not answering the set question etc (Figure 1). Students were asked to use their newly generated mark scheme to grade all of the photographs and to post this onto the discussion board. Once all participating students had completed this, the tutor provided the ‘course view’ together with an explanation of why we had given a photo a particular grading. We also provided a description of how this grade would relate to a piece of academic work on the programme (Figure 1), so linking the students back to the programme’s criterion referenced mark scheme. Further discussion and debate ensued on the online discussion forum, especially where there was discrepancy in views or students highlighted which aspects of the task they found particularly challenging.

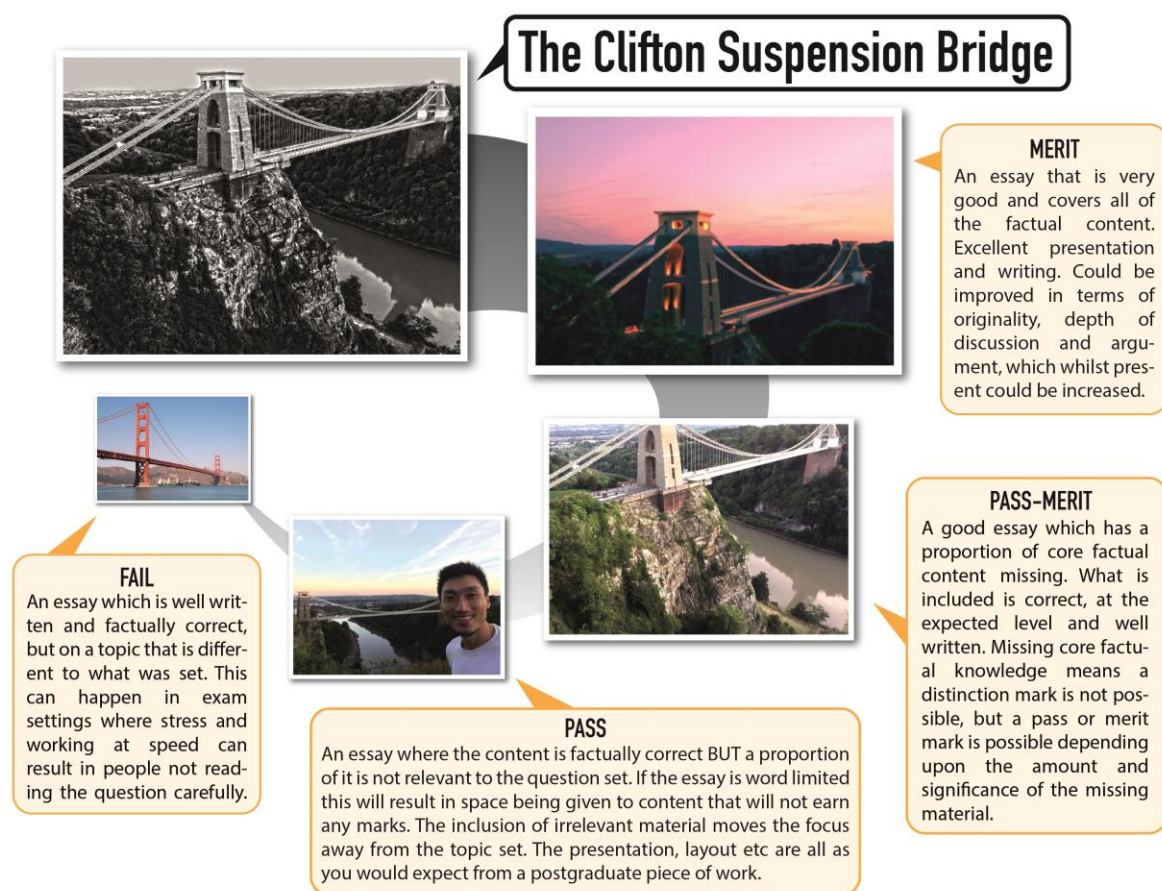


Figure 1. Illustration of photographs created for this task to illustrate common errors in written academic work.

Evaluation of the Assessment Literacy Task

Once this task was completed, students were asked to complete an online survey to provide their views on the task and its ability to help them understand criterion-referenced marking schemes. Some of these questions were free-text whilst others utilised a 5-point Likert Scale ranging from 1 'strongly disagree' through to 5 'strongly agree'.

Basic Linguistic Analysis.

In order to capture how our own mark scheme was structured we looked at the language used. This focused in particular on the use of adjectives. This basic linguistic analysis gave us a normative comparator for the schemes the students produced. To understand how the pattern of language used changed throughout our scheme we tallied the number of either adjectives or negative statements that appeared within each grade descriptor. This was repeated for every mark scheme created by the students. The highest number of adjectives in any individual scheme was used as a normaliser against which other tallies for each descriptor in that particular scheme were expressed. This enabled us to observe patterns in the use of both negative statements and adjectives in a single mark scheme and provided a way to visually compare language patterns between different schemes.

Evaluation of Impact on Summative Marks

The marks students had obtained for their first summative essay (1500 words) on the programme following the induction period when this marking scheme task ran were reviewed and considered in relation to i) whether the student had participated in the task and ii) how well their mark scheme compared linguistically to the programme's mark scheme.

Ethics

The University of Bristol's ethics committee confirmed that the online questionnaire was part of programme evaluation necessary to determine the usefulness of the task for students and identify ways in which it could be improved for future years.

Results

Student Demographics

18 students enrolled on this MSc programme in the 16/17 academic year. Whilst all students were invited to participate in this task only 11 students (61%) actively contributed. 22% (4/18) of the cohort were classified as overseas in terms of fee status and of these only 1 participated in this task. Of the 11 students that participated in the task 73% (8/11) completed the post task evaluation questionnaire.

Student's Evaluation of the Task

Of the eight students who completed the evaluation of the marking scheme literacy task, seven (88%) were aware of marking schemes and how these were used by markers. Their views on why such schemes were used focused on fairness and quality assurance e.g:

'To establish clear boundaries on standards of work.'

and

'To ensure fair marks across the board.'

Interestingly, the majority (6/8) of comments implied that marking schemes were only for markers with only two responses highlighting the benefits students can gain from understanding and using mark schemes:

So you know what to include in your work to get your grades.

Surprisingly 25% (2/8) of the group felt that mark schemes should only be made available to students after the assignment, a practice that would clearly not aid student's understanding of the criteria they will be assessed on. As this case study involved postgraduate students, it highlights the importance of ensuring staff involved in the delivery of Masters level education make no assumptions that their student's already have a good level of competency and familiarity with mark schemes and how they can be of benefit.

75% (6/8) of the students completing the post task evaluation either agreed or strongly agreed that the task had given them a better understanding of mark schemes. The other 2 students neither agreed nor disagreed. 63% (5/8) of the students agreed or strongly agreed that they would be able to use marking schemes to grade their own work after completing this task. When students were asked what they had learnt from the mark scheme task, responses ranged from understanding what is needed to obtain top marks to appreciation of how difficult writing mark schemes could be. For example:

Better understanding of grade boundaries and what is required to achieve them.

Fully understand what the point of the assignment is before attempting to answer

compared to:

It is hard to make a marking scheme that accommodates every possible scenario.

75% (6/8) of students agreed or strongly agreed that the task was an enjoyable experience, with the remaining two students having a neutral view on the experience.

63% (5/8) of students agreed or strongly agreed that the task would help them to prepare for future assignments, supporting the use of tasks such as these for future cohorts.

Linguistic Analysis and Comparison of Mark Schemes

Basic linguistic analysis of the programme's criterion-referenced mark scheme revealed a pattern in the use of adjectives. Adjectives are used throughout the programme's mark scheme, but the volume of their use decreases as you pass from the top end of the scheme through to the lower end (Figure 2, labelled C). As you would expect the adjectives at the top end of the mark scheme record the extent of how well work has been done – *outstanding, exceptional, extensive, excellent* or the quality of the work – *clear, fluent, accurate, relevant, independent*. Interestingly, 2 students (Fig 2, students 1 and 2) mirrored this pattern of adjective use in their own criterion referenced mark schemes and 4 students (Fig 2, students 3, 4, 5 and 6) generated mark schemes that closely mirrored the pattern of the programme's scheme. The mark schemes generated by the other 5 students (Fig 2, students 7, 8, 9, 10 and 11) did not reflect this pattern of adjective use.

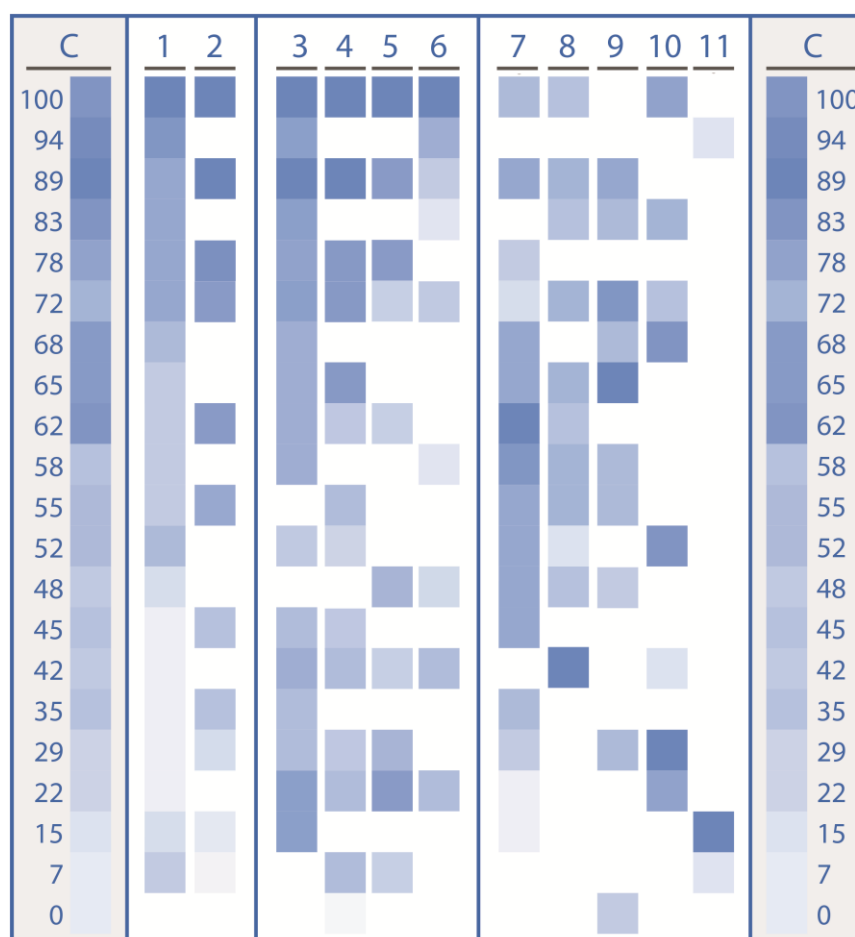


Figure 2. Diagrammatic illustration of simple linguistic analysis of the programme's mark scheme (C) compared to that of the 11 students who took part in the task.

Impact on Summative Work.

Whilst student feedback suggested that this task had helped individuals to understand the programme's mark scheme and feel more prepared for assessments, we were interested to see if this translated into any impact on their assessed written work. We considered the marks obtained from this group of students in their first written assignment, a 1500-word essay, which occurred within 3 months of the marking scheme task. Marks obtained for this task were normally distributed (Shapiro-Wilk test $p=0.1269$) and ranged from 20 to 76.5% with a median mark of 64% for the cohort ($n=18$). The marks obtained by those that did the marking scheme task were not significantly different to those that did not actively participate in this task ($63.7\pm 1.46\%$ vs. $58.5\pm 7.33\%$, unpaired T-test with Welch's Correction $p=0.399$, Fig 3a). The

difference in marks for the first essay also did not differ significantly for individuals that developed their own mark scheme in a linguistic manner similar to the programme’s scheme (n=6) compared to individuals (n=5) who’s own mark scheme did not mirror the style present in the programmes scheme (64.0±1.72% vs. 63.4±2.69%, unpaired T-test) Fig 3b). Whilst disappointing that a statistically significance impact on assessment was not illustrated it is important to note that the numbers of students in this study were small and hence it is not surprising that the results do not support those shown by Rust et al in 2003 where a student cohort of 300+ was used to show a positive impact of activities around mark schemes on subsequent academic work.

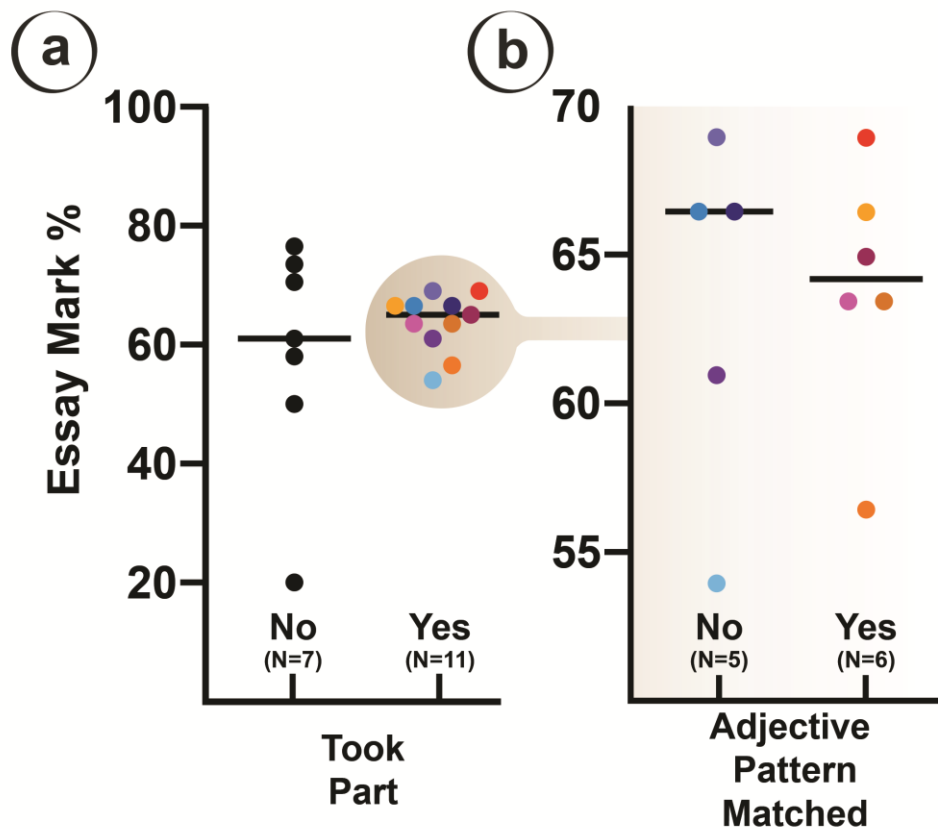


Figure 3a and 3b: Illustration of the range of marks obtained in the first summative essay following the assessment literacy task.

a) Range of marks by students who had (n=11) and had not (n=7) participated in the assessment literacy task.

b) Range of marks obtained by students who did participate in the assessment literacy task based on whether simple linguistic analysis of their own mark scheme matched (n=6) or not (n=5) with the programme’s mark scheme in terms of the pattern of the use of adjectives. Horizontal bars represent the median of each group.

Whilst no significant impact was illustrated on assessment marks it is worth highlighting that the range of marks obtained by those that did not engage in the task was significantly more variable than by those who had engaged (F-variance test, $p=0.0003$). Suggesting that this task may be able to help standardise and bring consistency to the performance and expectations of students who come from a range of academic backgrounds. Furthermore, the student cohort who were offered participation in this task reported very high levels of satisfaction with the programme's assessment practices when asked in an annual University wide survey. 100% (28% response rate) of students agreed that the marking criteria were clear in advance and that marking seemed fair. Whilst it is difficult to attribute any change in these scores directly to the introduction of this assessment literacy task this score of 100% in each parameter was an improvement over scores obtained in the previous year (93% and 86% respectively).

Discussion and conclusion

Results from the National Student Survey (NSS) have illustrated that a number of Higher Education institutions face difficulties in the area of assessment and feedback (Office for Students, 2019). This has resulted in many institutions considering how they can help students understand what is required from them in terms of assessments and helping them to understand how academics grade their work, i.e. a student's assessment literacy. As discussed previously, other authors have shown that interactive, supportive sessions are necessary to ensure students engage with tasks that are aimed at improving their assessment literacy and these can be difficult to deliver when students are remote and off-campus due to the need for tacit knowledge exchange. This case study illustrates that interactive campus-based tasks designed to improve student's assessment literacy can be converted into e-learning sessions that enable the explicit exploration of tacit knowledge.

Whilst no clear link between students completing this online task and an improvement of summative assessment marks was observed, it is important to note that the student cohort is small to observe such changes. Student feedback on the task was positive and suggested that the task has merit in terms of helping students feel confident about

how marking works and the programme's assessment practices and helping over 60% of the cohort feel more confident about preparing future assessment work. This positive impact on students' views of assessment practice is sufficient to support continued use of the task in future years. However, it is important to consider limitations of the task and whether modifications would be useful for future use. Some students indicated that writing a mark scheme on a topic they were not 'experts' in was challenging in terms of identifying suitable categories within the mark scheme. However, this provided the tutor with the opportunity to highlight the important link between intended learning outcomes and assessment. For example the tutor posed the question '*...would your scheme change if you were marking work on a course teaching photography compared to a course teaching bridge architecture?*' which allowed students to consider the importance of understanding the purpose of a task and realising that how a mark scheme is designed will depend upon what skills/knowledge/attributes the student is expected to display and thus direct students to the intended learning outcomes for the programme and its units. Within this discussion the issue of education level was also raised, ie the tutor asked '*what do you think would change about your scheme if you were teaching undergraduates as opposed to postgraduates?*' enabling the link to be made to study descriptors for Bachelor's as opposed to Master's level of study hence helping students to be aware of the need for critical analysis and originality in their work on this Master's programme.

Once students were using their schemes to mark the photos we had provided a wide level of marking variation was evident. This provided the tutor an opportunity to highlight key aspects of the quality assurance processes employed by the programme and hence build the student's confidence in the fairness of the marking they would receive.

Tutor: We ensure that all markers of your coursework use our single course mark scheme that the Unit leaders composed following discussion and review. This mark scheme is based on the 21point scheme in the University's taught code of practice that was drawn up from the QAA standards. It is also important to reassure you that all of your work is independently marked by two people. If there is significant variation, then the two markers meet to discuss their views

and try to decide upon a final mark. If they continue to disagree we will use a third marker.

The use of photos to mimic specific academic writing errors, such as plagiarism or not answering the question posed, appeared to resonate with students:

Student: I imagine this is similar to a student misreading a question and going off at a tangent in an essay. Ending up with a well written piece of work, but not actually answering the question posed.

Without prompting and within the task itself students could see the benefit of improving their assessment literacy e.g:

Student: It certainly made me more aware about what is expected of us in our written work.

and

Student: Overall I found it an interesting exercise and it definitely made me evaluate the content of a piece of work and how to breach grade boundaries.

One concern is the failure of all students to engage with the task. There will be many reasons for this, ranging from individuals who already feel confident in their understanding of mark schemes through to those that do not appreciate the relevance or potential benefits or those that feel comfortable participating in tasks that are outside their specialist knowledge. In the future it will be worth considering ways to positively encourage all students to engage and participate. Making it compulsory may not be the best approach as it is during the induction period and students have not bonded as a cohort and some are anxious about communicating in an unfamiliar online forum on an unfamiliar topic. Unfortunately, we are not able to determine how many students participated by observation whilst not actively posting comments, although work by Rust et al (2003) suggests that active participation aids the positive benefits of such tasks. It would also be interesting to compare delivery of this task online, both in terms of staff time and student benefit, to that when its delivery in a face to face setting.

References

- Anderson, J & Ostman, L. (2015) A transactional way of analysing the learning of “tacit knowledge”. *Interchange*, 46(3), 271-287.
- Asher, D & Popper, M. (2019) Tacit knowledge as a multilayer phenomenon: the “onion” model. *Learning Organization*, 26(3), 264-275.
- Higher Education Statistics Agency (HESA) (2019). *Who’s studying in HE?* Viewed November, 26th, 2019. <https://www.hesa.ac.uk/data-and-analysis/students/whos-in-he>.
- Lucas, L & Webster., F. (1998). Maintaining standards in higher education? A case study. In D.Jary, & M. Parker (Eds.), *The New Higher Education: issues and directions for a Post-Dearing university* Staffordshire University Press.
- O’Donovan, B., Price, M., & Rust, C. (2001) The student experience of the introduction of a common criteria assessment grid across an academic department. *Innovations in Education and Teaching International*, 38(1), 74-85.
- 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.
- Panahi, S., Watson, J., & Partright H. (2012) Social Media and Tacit Knowledge Sharing: Developing a Conceptual Model. In *World Academy of Science, Engineering and Technology*. Paris, France: World Academy of Science, Engineering and Technology (WASET), 1095-1102.
- 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.
- Rust, C., Price, M., & O’Donovan, B. (2003), Improving Students’ Learning by Developing their Understanding of Assessment Criteria and Processes. *Assessment and Evaluation in Higher Education*, 28(2), 147-164.
- Salmon, G. (2005), Flying not flapping: a strategic framework for elearning and pedagogical innovation in higher education institutions. *ALT-J, Research in Learning Technology*, 13(3), 201-218.
- 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 and Evaluation in Higher Education*, 38(1), 44-60.
- Ubon A., & Kimble C. (2002), Knowledge management in online distance education. *Proceedings of Networked Learning Conference*. Viewed November, 26th, 2019. <http://www.networkedlearningconference.org.uk/past/nlc2002/proceedings/papers/26.htm>.
- Webster, F., Pepper, D., & Jenkins, A. (2000). Assessing the undergraduate dissertation. *Assessment and Evaluation in Higher Education*, 25(1), 72-80.