The road to self-assessment: exemplar marking before peer review develops first-year students' capacity to judge the quality of a scientific report

Robyn Yucel[a]*, Fiona L. Bird[b], Jodie Young[c] and Tania Blanksby[d]

[a]Student Learning, La Trobe Learning and Teaching, La Trobe University, Melbourne, Australia; [b]Department of Zoology, La Trobe University, Melbourne, Australia; [c]Department of Genetics, La Trobe University, Melbourne, Australia; [d]Faculty of Science, Technology and Engineering, La Trobe University, Melbourne, Australia

Lack of clarity about assessment criteria and standards is a source of anxiety for many first-year university students. The Developing Understanding of Assessment for Learning (DUAL) programme was designed as a staged approach to gradually familiarise students with expectations, and to provide opportunities for the development of the skills required to successfully complete assessment tasks. This paper investigated the students' perceptions of the first two components of the DUAL programme, which assist first-year biology students to engage with stated assessment criteria and standards in order to develop their capacity to make judgements about scientific report exemplars, their peers' scientific reports and ultimately their own. The study found strong evidence (96% of responses) that the marking and discussion of exemplar reports with peers and demonstrators clarified expectations of scientific report writing. A key feature of this element of DUAL was the opportunity for structured discussion about assessment criteria and standards between peers and markers (demonstrators). During these discussions, students can clarify explicit statements and develop a tacit knowledge base to enhance their ability to judge the quality of others' work and their own. The peer review exercise (the second element of DUAL) was not rated as highly, with 65% of students finding the process helpful for improving their report. The negative reactions by a sizeable minority of students highlight the need to clearly communicate the expectations and benefits of peer review, with a focus on how the process of *giving* feedback to peers might benefit a student as much as *receiving* feedback on their own report.

Keywords: assessment criteria; formative assessment; peer feedback; peer review; scientific writing

Introduction

First-year experience programmes have greatly improved first-year students' social transition to university by providing opportunities to enhance their sense of belonging and integration into the university community (Kift, Nelson, and Clarke 2010). However, until recently, little attention has been paid to the role of assessment in assisting students during transition (Nelson et al. 2011). Assessments in the first year can be a source of intimidation and confusion (Calder and Daly 2009), as large classes of students are often assessed with tasks which are scheduled late in the

---

*Corresponding author. Email: r.yucel@latrobe.edu.au

semester, are heavily weighted towards summative assessment and do not provide students with feedback or encourage engagement (Krause and Coates 2008). Kift and Moody (2009) argue that the key to supporting student learning, success and retention lies with well-designed assessment and feedback practices.

One of the biggest sources of anxiety for first-year students around assessment is knowing what is expected. Most first-year students are unfamiliar with tertiary assessment practices, relating to both the types of assessment tasks they will complete and the academic skills they will require to complete them (Kift and Moody 2009). Even if students have attempted a similar type of assessment task (e.g. a laboratory report) in secondary school, the format, style of writing, referencing system and academic standards expected may all be different at university. Clarifying expectations with clearly stated goals, criteria and expected standards is critical to good curriculum design (Nicol and Macfarlane-Dick 2006). The provision of marking rubrics and exemplars, which detail the assessment criteria and standards, can clearly show students the goals to aim for (Hattie and Timperley 2007). However, unless students actively engage with them during the assessment process, rubrics and exemplars fail to meaningfully transfer knowledge of standards to students (O'Donovan, Price, and Rust 2004; Rust, O'Donovan, and Price 2005; Nicol and Macfarlane-Dick 2006).

Prior to writing their first assessment piece, students who engage with a rubric, by marking an exemplar report against the criteria and descriptors of standards, develop their understanding of what is expected (Rust, Price, and O'Donovan 2003). Knowledge of expectations and standards can be further developed in students through participation in post-exemplar marking moderation discussions of the assessment criteria and exemplars (Rust, O'Donovan, and Price 2005). Subsequent to that, feedback on drafts of written work prior to submission for summative assessment can further reduce first-year students' assessment anxiety. If this feedback is given by peers, it gives all students a further opportunity to actively engage with the assessment process (Boud 1995).

Peer assessment is likely to be an unfamiliar process for most first-year students (Ballantyne, Hughes, and Mylonas 2002), and studies have found that university students have mixed reactions to it. Reported perceived benefits include improved ability to make judgements about others' work and their own (Topping 1998), learning from others (Ballantyne, Hughes, and Mylonas 2002; Venables and Summit 2003), developing skills for future employment (Ballantyne, Hughes, and Mylonas 2002) and developing critical thinking and analytical skills (Vickerman 2009). However, some studies report that students are uncomfortable with peer assessment. Major criticisms include a perception of bias due to friendships within student cohorts (Magin 2001; Ballantyne, Hughes, and Mylonas 2002; Liu and Carless 2006; Papinczak, Young, and Groves 2007; Patton 2011), students not taking peer assessment seriously (Papinczak, Young, and Groves 2007; Patton 2011) and lack of confidence in their own and their peers' ability to provide feedback (Ballantyne, Hughes, and Mylonas 2002; Smith, Cooper, and Lancaster 2002). While the term 'peer review' outside of the higher education context is taken to mean a summative assessment of a paper's suitability for publication, we use the term 'peer review' in the university classroom context to mean formative assessment given by peers before submission of an assessment task for summative assessment (Patton 2011; Mulder et al. 2013; Nicol, Thomson, and Breslin 2014).

Students perceive formative peer review more positively than summative peer assessment (Vickerman 2009; van Zundert, Sluijsmans, and van Merriënboer 2010). Peer review is considered to be a 'safer' and less-threatening exercise than summative peer assessment, particularly at first-year level (O'Donovan, Price, and Rust 2008; Vickerman 2009), and to have a greater positive impact on student learning (Liu and Carless 2006).

Summative peer assessment is becoming more prevalent in higher education curricula (Vickerman 2009; van Zundert, Sluijsmans, and van Merriënboer 2010). However, it needs to be recognised that the ability to make judgements about the quality of work for assessment is a developmental skill. Therefore, a staged approach, starting with formative peer review in the first year, can help to lower student anxiety about the process and build students' skills (Smith, Cooper, and Lancaster 2002; Venables and Summit 2003). It has been shown that, with adequate training in giving and receiving feedback, and with discussion of what constitutes quality work, students are capable of providing useful feedback to peers (McMahon 2010; MacDonald 2011; Rubin and Turner 2012).

A key to successfully training students in peer review and assessment is to provide opportunities for discussion of assessment criteria and standards (Sadler 2010). During discussion with peers and tutors/demonstrators, explicit statements about assessment criteria can be analysed and knowledge about them shared (O'Donovan, Price, and Rust 2008). Discussion also provides an opportunity for developing the ability to make qualitative judgements about assessment standards, by developing an individual's tacit knowledge base through the process of discussing judgements and comparing them with others (Sadler 2010; Bird and Yucel 2013). The process of exemplar marking and peer review is essentially the same, with the students moving from reviewing exemplars to reviewing their peer's work. We have developed a comprehensive assessment programme which includes the use of exemplar marking and group discussion as an effective preliminary activity for peer review.

The programme: Developing Understanding of Assessment for Learning (DUAL)

The DUAL programme was designed with the dual purpose of improving marking reliability in an inexperienced teaching team and developing scientific writing skills

Table 1. Timeline of student activities within the Developing Understanding of Assessment for Learning (DUAL) programme across weeks 8–12 of semester.

Week	Student learning activity
8	Students assess two exemplar reports using the rubric and participate in moderation discussion with peers and demonstrators Annotated exemplar reports (exemplars) are made available on subject webpage Students draft their first report
9	Students participate in a peer review of report one and improve their draft in response to feedback
10	Students submit report one for assessment
11	Report one returned Students complete an action plan which outlines how they plan to address feedback from markers
12	Students submit report two with action plan

in a large first year biology student cohort. The DUAL programme includes a staged training programme for the team of laboratory demonstrators (evaluated in Bird and Yucel 2013), alongside a series of classroom activities which build skills and knowledge in students over the semester (Table 1). The underlying principle of the programme is the provision of opportunities for discussion of criteria and standards.

The DUAL programme for students has four specific aims:

(1) To clarify report-writing conventions and standards through an exemplar report marking and moderation procedure;
(2) To improve the quality of students' reports through a process of peer review before the final submission;
(3) To improve the quality and quantity of feedback to students through the markers' use of a feedback code to identify errors in students' reports; and
(4) To improve student engagement with feedback through the use of a feedback code and an action plan, which asks students to use feedback on their first report to identify areas to improve in their second report.

Of the four components of the DUAL programme for students outlined above, the first two are the focus of this paper.

Exemplar report marking and moderation activity

Two exemplar first-year laboratory reports were produced by a research assistant. A student report on a discontinued topic was manipulated, so that one version provided an example of a very high standard (assessed as 94% by two of the authors), and the other provided an example of an average standard (assessed as 72.5% by two of the authors). Students were given the two reports in a laboratory class, and asked to mark them according to a standard rubric used to assess laboratory reports in first-year biology. Then, in groups, students discussed the marks they had given, justifying the marks they had awarded to particular sections. After the moderation activity, the demonstrators led a discussion during which students were told the marks awarded to each report and were shown the reports with an annotated feedback code developed by one of the authors.

Peer review exercise

One week before final submission of their first report, students brought a draft report to the laboratory class. The reports were collected and re-distributed to students by the demonstrators. Each student was asked to mark another student's report anonymously, using the marking rubric. Students could use this feedback to edit their reports before final submission the following week. During the class, students were encouraged to seek clarification from their demonstrators about any feedback that was confusing or thought to be erroneous.

The study

The purpose of this study was to explore the impact of the DUAL programme on students' report marks and perception of their own learning in terms of: (a) their

understanding of the task requirements and (b) the perceived effect of the DUAL programme on their report-writing skills.

Report marks

Participants

Students enrolled in the subject Animal Evolution and Diversity were the focus of this study, and a comparison was made between cohorts in the years 2009 (pre-DUAL: 373 students) and 2010 (post-DUAL: 403 students). The median Australian Tertiary Admissions Rank (ATAR) score was calculated to confirm that the cohorts of students were of similar academic potential (67.6 and 68.5 in 2009 and 2010, respectively). ATAR scores range between 0 and 100.

Comparison of report marks

Students submit two reports from two different investigations in the subject for marking by the laboratory demonstrators. Average laboratory report marks for both reports were compared between student cohorts (2009 and 2010) using a Student's *t*-test, to investigate if the DUAL programme had improved performance on laboratory report-writing tasks. Marks awarded to reports one and two were analysed separately. All marks were arcsine transformed prior to analysis to improve the normality of the data-set. All data were checked for homogeneity of variances before analysis, and the assumption was met in all cases.

Student perceptions of DUAL

To further explore the value of the DUAL programme, the 2011 cohort was surveyed to gain an insight into the reasons why students found aspects of DUAL helpful or unhelpful. A complete or partially complete questionnaire was returned by 369 students out of a total of 488 students enrolled. The questionnaire comprised five open-ended questions, which sought to elicit students' perceptions of the effect of the four components of the DUAL programme on their own learning. This paper focuses on the first two components of the DUAL programme, the exemplar marking and discussion activity and the peer review exercise, which were surveyed in questions one and two:

(1) Did the exemplar report marking exercise and group discussion of standards help you to understand what was expected in a first-year laboratory report? If yes, how? If no, why not?
(2) Did participating in the peer review exercise help you to improve your report before submission? If yes, how? If no, why not?

The answers to these questions were classified into one of four categories: *yes*, *no*, *ambivalent* and *did not participate*, and the responses for the whole cohort were quantified. Because the reasons students gave for each of their responses followed easily recognisable patterns, content analysis (Weber 1990) was used to identify these emerging themes. The four authors of the paper agreed upon a limited number of categories which captured the range of student responses for the exemplar report

marking discussion and peer review activities. All responses for each question were then coded by an individual author using the agreed categories. Some students' comments related to more than one category; therefore, the sum of student responses for all categories is greater than the total number of responses. The frequency of responses for each category was recorded. Selected quotations are presented to illustrate major emergent themes.

Results

Comparison of student report marks pre- and post-DUAL

Contrary to our expectations, students who participated in DUAL achieved significantly lower average marks for both reports one (*t*-test, $t = 3.84$, d.f. $= 765$, $p < 0.001$) and two (*t*-test, $t = 2.43$, d.f. $= 778$, $p = 0.015$) than students who did not participate in DUAL. Although average marks differed significantly between years, the variation was minimal (report one: $73 + 13\%$ pre-DUAL compared to $69 + 13\%$ post-DUAL, and report two pre-DUAL $73 + 14\%$ compared to $71 + 13\%$ post-DUAL). Interestingly, an additional 10% of the class improved their mark from report one to report two after they participated in DUAL (64%), compared to 54% of students who improved their mark from report one to two without the DUAL programme.

Exemplar report marking exercise

In question one of the student survey, students were asked if the exemplar report marking exercise and group discussion helped them to understand what was expected in a first-year laboratory report. Of the 322 responses that were returned, the overwhelming majority of students (96%) agreed that the exemplar report marking and group discussion helped in their understanding of the expectations. Three per cent of students reported that the exercise was not helpful, and the remaining 1% were either ambivalent about the exercise or absent for the exercise (see Table 2).

The majority of students (75% of 'yes' responses) reported that the marking and discussion of exemplar reports was helpful because it clarified expectations of scientific report writing. Typical comments included:

> Yes very much so. I would not have known what to do if we had not done this. It gave me an accurate expectation.

Table 2. Reasons given by students who found the marking and discussion of exemplar reports helpful. Where a response falls into more than one theme, it has been counted under each of the relevant themes.

Category	No. of students
Clarified expectations generally	240
Clarified expected standard	34
Raised error awareness	16
Discussion of reports was of value	8
Integrated with other subject resources	4
Staff helped me	2
No reason/not applicable to question	40

There was also a sense in some responses that the exercise helped to reduce anxiety around report writing, with one student commenting:

> Yes, otherwise I would've panicked and had no idea what to do.

In most cases, students did not identify whether they were clearer about expected criteria or standards for the report. However, 11% of students who responded 'yes' to question 1 specifically mentioned clarification of standards as being helpful, such as in this comment:

> It was good to get familiar with the criteria that would be used to mark our reports as senior school prac reports are a lot more basic in structure than uni reports. It also gave us a rough idea of the standards we would need to write at to achieve a decent mark.

A small proportion of students (5% of 'yes' responses) commented that looking at the two exemplar reports raised their awareness of potential errors to avoid, for example:

> Yes, the exemplar report gave me a template to work from and I was able to use the corrections we made on it to ensure I didn't make those mistakes on my own.

Interestingly, only one student commented that the exercise had helped develop self-assessment skills:

> Yes, you can use the values to grade your own prac report.

Peer review exercise

In question two of the student survey, students were asked if the peer review exercise helped them to improve their report before they submitted it for assessment the following week. Of the 313 responses returned, the majority of students (65%) agreed that the peer review process helped them to improve their scientific report before submission. There were 18% of students who disagreed that the peer review exercise was helpful, 8% were ambivalent about its helpfulness and 9% indicated that they were absent on the day the peer review exercise was conducted.

'Yes' responses

The responses of students who felt the peer review exercise was helpful ($N = 203$) were coded and eight categories were identified (see Table 3).

Of those who responded positively, more than half (58%) found the peer review exercise helpful because it highlighted omissions or errors and areas for improvement, for example:

> Yes. My peers picked up on mistakes which I was not aware of which further deepened my understanding.

> Yes! It showed areas of improvement which would assist in further consolidating my work before submission.

Table 3. Reasons given by students who found the peer review exercise helpful. Where a response falls into more than one theme, it has been counted under each of the relevant themes.

Category	No. of students
Highlighted omissions or errors	64
Highlighted areas for improvement	54
Resulted in improvement generally	43
Learned from others	42
Clarified expectations	24
Discussion was useful	5
Time management	2
No reason given	19

Twenty-one per cent of students who responded 'yes' to question 2 reported that participating in the peer review process helped them to improve the quality of their report. Typical comments included:

> Yes, I found that the person who marked my report gave me feedback that definitely helped to improve my report before I handed it in to my demonstrator.

A further 21% of students who responded 'yes' to question 2 commented that they could learn from both others' errors and aspects others had done well:

> Yes! It helped me realise what is relevant info. It was beneficial and nice to have another science student read my report and provide feedback. It was beneficial to read someone else's and put it in perspective to mine.

A smaller number of students (12% of 'yes' responses) felt that the criteria and/or standards of report writing were clarified by the peer review exercise, for example:

> Yes, because it showed me what is needed to include and gave me an idea of the standard.

A comment that the discussion with other students helped to clarify criteria or errors was made by 2% of students who responded 'yes' to question 2, for example:

> Yes certainly as we can discuss our points of view and learn new techniques with our peers.

> 'No' responses

The responses of students who felt the peer review exercise was not helpful ($N = 56$) were coded and five categories were identified (see Table 4).

Sixty-one per cent of students who did not find the peer review exercise helpful were unsatisfied with either the quality or quantity of feedback given, as in this comment:

> No not at all. If anything it made it worse. I was told I did really well and when I got marks back I didn't do that well. Get rid of this exercise! Maybe let people read other

Table 4.   Reasons given by students who considered their reports were not improved by the peer review exercise. Where a response falls into more than one theme, it has been counted under each of the relevant themes.

Category	No. of students
Poor quality feedback	18
Insufficient feedback	16
Peer inexperienced	17
Peer not interested	4
No reason given	9

> reports but no mark for it give them false feedback! I was gonna re-write report but due to good peer review decided not to.

A further 30% of students who responded 'no' to question 2 identified the reason for poor feedback from peers being the inexperience of the peer markers:

> No. Because fellow students are also in first year and are also doing it for the first time. They do not necessarily understand the requirements fully yet either, and the comments can be very far from the mark.

A small proportion of students (7% of 'no' responses) identified lack of interest of the peer marker as a reason for poor feedback, for example:

> No. Feedback was very minimal, incorrect and not useful. Few students seemed to take it seriously.

Discussion

Evaluation of DUAL using students' marks

There was no improvement in students' marks after the DUAL programme was introduced in 2010 when compared with the 2009 cohort. One possible reason for this was that because of the additional activities associated with the DUAL programme: report one was due in week 10 of the 12-week semester compared to week 9 in 2009, bringing the deadline closer to the end of semester when students start to focus on examination preparation. Careful timetabling of assessment tasks across a semester, particularly in the first semester of first year when students are still adjusting to university life, is critical for a manageable student workload (Kift and Moody 2009).

The lack of improvement in student performance may also have been due to the way the reports were marked, rather than the actual quality of the reports submitted by students. Bird and Yucel (2013) quantified the effect of the DUAL programme on marking reliability of demonstrators, and showed that, after participation in DUAL, demonstrators marked reports significantly closer to the marks awarded by an expert marker of the discipline. The demonstrator training section of DUAL included an exemplar report marking and moderation discussion exercise, which supported explicit knowledge exchange about marking between staff and the development of tacit knowledge of standards in individual markers (Bird and Yucel 2013). Demonstrators commented that the specific nature of the marking rubric,

which describes the assessment criteria and standards, made it easier to mark reports. However, they also reported that their expectations regarding the quality of student reports were raised, as students should now have been clearer about the expected criteria and standards. Thus, the DUAL programme, whilst developing expertise in marking and understanding of standards by individual markers, also influenced their approach to marking and therefore the marks awarded to the reports in 2010.

A further possible reason why this study was unable to detect an improvement in marks between the 2009 and 2010 cohorts was that it is not possible to control for the myriad variables that affect the quality of a student's written work on a given occasion. These variables may include: motivation levels; external demand on time from family, employment and social groups; workloads from other subjects; variation in the marking of demonstrators; and changes in the levels of ability in different cohorts of students. Confining evaluation of the efficacy of curriculum change to marks alone also takes a narrow view of learning, which excludes the learner's perspective from the evaluation. Perhaps more important than numerical values are the confidence and reduction in anxiety that come from being clear about assessment criteria and standards, particularly in the first year. Thus, we contend that the qualitative data about student perceptions of the DUAL programme are more revealing about the benefits to students.

Evaluation of DUAL from the student perspective

Marking and discussion of exemplar reports

In the student questionnaire, the marking of exemplar reports and subsequent discussion of these with peers and demonstrators was reported to clarify expectations of scientific report writing by the overwhelming majority (96%) of the students. A rubric which gives students guidance on both criteria and standards is a first step towards clarifying assessment requirements for students. However, simply providing a rubric is insufficient to foster meaningful student engagement with the expected criteria and standards. Studies have consistently shown that the key to clarifying assessment criteria and standards for students is to ensure students actively engage with them (e.g. Rust 2001; Rust, Price, and O'Donovan 2003; Bloxham 2009). The use of exemplars has been widely advocated as an effective means to communicate expected standards to students (Ballantyne, Hughes, and Mylonas 2002; Orsmond, Merry, and Reiling 2002; Rust, Price, and O'Donovan 2003). The extremely positive evaluation of the exemplar report marking exercise in the DUAL study provides strong evidence for the value of including this exercise in a first-year assessment programme.

Peer review exercise

While feedback on the peer review exercise was generally positive, it was not rated as highly as other aspects of the DUAL programme, with only 65% of students finding the process helpful for improving their report. Similar results were found in a study by Vickerman (2009), where 59% of students agreed or strongly agreed that formative peer assessment of literature reviews had helped them to better understand the assessment process. Students in the DUAL study who found the peer review exercise helpful mostly reported that the exercise pointed out what they had done

wrong or showed them where they could improve. Of the students who did not find the peer review process helpful, the majority reported that they received either insufficient or poor quality feedback. Some students identified the inexperience of the peer markers as a cause of their dissatisfaction with the feedback they received, and this issue has been reported elsewhere in the literature (Ballantyne, Hughes, and Mylonas 2002; Patton 2011). A smaller number of students in the present study reported that some peers did not engage with the process of peer review, a finding that has also been reported by Cartney (2010). The higher education assessment literature reports mixed evaluations of peer review and peer assessment. In general, summative peer assessment, where students' grades are based on a proportion of marks given by peers, is regarded less favourably by students than formative peer review, where peer marks are not counted towards final grades (Patton 2011).

Theorising the practice: how exemplar marking and peer review builds knowledge of scientific report writing

The DUAL programme provides a temporary community of practice (Lave and Wenger 1991) for students to engage in a structured discussion with both peers and markers about the two exemplar reports and peer's reports in relation to the criteria and standards in the rubric. In this community of practice, knowledge is shared amongst novices and experts as students articulate their own understanding of the criteria in the rubrics (thus forcing them to be clear about and justify their views), and have an opportunity to compare their understanding with that of others. This approach is underpinned by the recognition that learning is not simple transmission or transferral of information from teacher to student. Making clear and explicit statements in the rubric does not ensure that students fully grasp the expected criteria and standards, as this assumes a simple one-way transmission of knowledge and ignores the complex nature of the learner, who is not an empty vessel waiting to be filled with knowledge (Freire 2000), but rather a complex being who brings previous conceptions and experience to the learning process. Learners must construct new knowledge by building a conceptual bridge between their existing knowledge and the target concept (von Glasersfeld 2007). This construction of knowledge involves the development of both *explicit* and *tacit* knowledge.

Two types of judgements are involved when students are comparing the exemplars with the rubric. Statements in the rubric contain interwoven aspects related to both criteria and standards. For example, our rubric includes a statement about the abstract in the satisfactory band: 'Abstract clearly states overall aim of investigation, hypothesis/es to be tested and includes a concise and correct summary of the major findings'. This statement has two different types of expectation. First, it describes what must be present to be a complete abstract (the criteria), but it also describes the *quality* of what is present (the standard). If an aim is present, the criterion of 'states the overall aim of the investigation' is satisfied, but whether it does this *clearly* is related to the quality of the statement of the overall aim (i.e. the standard). Similarly, 'concise and correct' are quality judgements that relate to the expected standard. Understanding of the *criteria* largely involves the unpacking of explicit statements that can be clarified relatively straightforwardly during discussion. Students need to match criterion statements in the rubric (e.g. the abstract states the aim and hypothesis/es) with instances of those items in the exemplars. This requires that students both understand what the statement means and are able to apply that understanding

in order to recognise an example of it in the exemplar. The discussion provides an opportunity for students to share knowledge by helping each other to clarify the meaning of the explicit statements in the rubric.

An understanding of the *standards* requires a component of judgement, and thus also involves a tacit dimension (Polanyi 1958). Making a judgement involves applying understanding of the quality of several particulars simultaneously to the quality of something as a whole. According to Polanyi, it is not possible to be aware of all of these particulars in the process of making more holistic judgements about quality, and thus the particulars remain tacit during the act of judgement (Polanyi 1962). There is some debate in the higher education assessment literature about the extent to which the tacit knowledge involved in making quality judgements can be shared (see Bird and Yucel 2013). While some authors (e.g. Rust, Price, and O'Donovan 2003; O'Donovan, Price, and Rust 2008) argue that, during discussions about applying assessment criteria and standards, both explicit and tacit knowledge are shared, this view appears to contradict Polanyi's original (1962) conception of tacit knowledge as being unable to be articulated, and therefore unable to be shared (Gourlay 2006a, 2006b; Klein 2008). What seems more likely to happen is that, through the discussion, students build their own tacit knowledge base through the sharing and clarification of explicit knowledge about criteria and standards, so that, when they make a quality judgement, this knowledge is used tacitly in the act of judgement. Students also receive feedback on that judgement from others, including demonstrators, which can help them to assess their own abilities to judge the quality of a scientific report, and thus ultimately improve their ability to self-assess their own work.

A recognition of the role of both tacit and explicit knowledge in building students' skills in evaluating written reports, and eventually evaluating their own, emphasises the central importance of including discussion of exemplars in the students' assessment programme. Although more able students may be capable of developing this expertise independently, scaffolding student learning in this way gives all students a greater opportunity to develop a level of connoisseurship (Eisner 1985) when judging how well their own work exemplifies the expected criteria and standards of a scientific report. It would be interesting to more rigorously investigate the effect of discussion of marking exemplars on building students' expertise in making tacit quality judgements about assessment standards. With the proliferation of online and blended learning modes in higher education, it would also be useful to investigate any differential effects in face-to-face vs. asynchronous online discussion settings.

Implications for practice

It is clear from student responses in the DUAL study that some students reacted quite negatively to being reviewed by their peers. When asked if the peer review process helped to improve their report, one student commented: 'No. The peer who corrected mine was an idiot and told me to correct things that were clearly wrong. Maybe if we ensured my peers had a brain it would have been useful'. Including a self-review activity before the peer review may help to build students' skills in applying criteria and standards, and help them to make a more explicit comparison between their own judgement of their work and the judgement of others. The perception of poor quality feedback from peers could also be partially overcome by

using peer review triads (Ballantyne, Hughes, and Mylonas 2002), where students receive feedback from two other peer reviewers rather than just one. For large classes, administering anonymous peer reviews in groups of three or more may present challenges, so it is encouraging that online peer review systems such as Calibrated Peer Review (Robinson 2001) and PRAZE (Mulder and Pearce 2007) have been developed that allow students to receive peer feedback from multiple peer reviewers with minimal administrative burden.

The disappointment expressed by some students with the peer review exercise highlights the need to clearly articulate the rationale and expectations for peer review prior to the exercise in order to give students more realistic expectations. Some students in the DUAL study seemed to perceive the peer review exercise as a sub-standard form of demonstrator feedback, with one student commenting:

> This was a really bad exercise. I felt that I didn't get the right feedback. If my demonstrator would have given me the feedback instead that would have been better. I put in a lot of work in my first draft and I wanted an experienced person to look at it and give me feedback! Bit disappointed!

The belief that the sole purpose of peer review is to receive high-quality feedback may lead to disappointment when it does not. Thus, students need to be made aware of the range of benefits they might obtain from the peer review exercise.

No student in the current study explicitly commented on how the process of *giving* feedback to peers might have improved their own ability to judge the quality of a scientific report, and the benefit this might have for their self-assessment skills for judging the quality of their own writing in future. Boud (1995) argues that students' active engagement in evaluating others work and evaluating the feedback they receive develops skills in self-assessment, making them more autonomous learners. Li, Liu, and Steckelberg (2010) found that students in their study benefitted more from the process of reviewing another student's work than from the feedback they received from their peers. Making this point clear to students may help to ameliorate some of the frustration experienced during peer review, as well as help them develop the metacognitive skills needed to become autonomous life-long learners. A practical strategy to emphasise to the students the value of *giving* feedback would be to awards marks to peer reviewers for the quality of their peer review (Ballantyne, Hughes, and Mylonas 2002). Bloxham and West (2004) found that awarding 25% of marks for a poster task to the quality of the peer marking helped students to recognise the benefits of giving peer feedback.

The present findings confirm that an exemplar report marking exercise followed by a discussion of the quality and features of the reports between peers, and peers and markers, is a valuable step in the development of students' knowledge and understanding of the assessment criteria and standards expected for the task. We argue that this type of activity should be included each time a new type of assessment task is introduced within a course of study. During discussions, students can clarify explicit statements, and develop a tacit knowledge base to enhance their ability to judge the quality of others' work and their own. First-year undergraduate students, in particular, will benefit from these scaffolding exercises, as they are required to tackle new and unfamiliar types of assessment. Providing students with opportunities to give and receive feedback on a piece of assessable work is common, but the findings of this study highlight the importance of clarifying the expectations

of exemplar marking and peer review to students. The study also highlights the need to emphasise the potential benefits of a peer review exercise, with a focus on how the process of *giving* feedback to peers might benefit a student just as much as *receiving* feedback on their own report. More broadly, while the rationale for new curriculum initiatives may be clear to those designing and teaching the curriculum, they need to be done 'with' the students, not 'to' them, by explicitly communicating the rationale behind tasks aimed to improve student learning.

Notes on contributors

Robyn Yucel is a lecturer in the Faculty of Science, Technology and Engineering's Academic Language and Learning Unit at La Trobe University, Melbourne. Her current research interests include assessment in higher education, academic language and learning in higher education science, and the nature of science in undergraduate science degrees.

Fiona L. Bird is a senior lecturer and head of the Department of Zoology at La Trobe University, Melbourne. Her current research interest is in the area of assessment in higher education.

Jodie Young is a lecturer in the Department of Genetics at La Trobe University, Melbourne. Her current research interest is in the area of assessment.

Tania Blanksby is a first-year coordinator of the Faculty of Science, Technology and Engineering at La Trobe University, Melbourne. Her current research interests include first-year experience and transition.

References

Ballantyne, R., K. Hughes, and A. Mylonas. 2002. "Developing Procedures for Implementing Peer Assessment in Large Classes Using an Action Research Process." *Assessment & Evaluation in Higher Education* 27 (5): 427–441.

Bird, F., and R. Yucel. 2013. "Improving Marking Reliability of Scientific Writing with the Developing Understanding of Assessment for Learning Programme." *Assessment & Evaluation in Higher Education* 38 (5): 536–553.

Bloxham, S. 2009. "Marking and Moderation in the UK: False Assumptions and Wasted Resources." *Assessment & Evaluation in Higher Education* 34 (2): 209–220.

Bloxham, S., and A. West. 2004. "Understanding the Rules of the Game: Marking Peer Assessment as a Medium for Developing Students' Conceptions of Assessment." *Assessment & Evaluation in Higher Education* 29 (6): 721–733.

Boud, D. 1995. *Enhancing Learning through Self Assessment*. London: Kogan Page.

Calder, A., and B. Daly. 2009. "Learning from Our Students." *Journal of Academic Language and Learning* 3 (1): A26–A37.

Cartney, P. 2010. "Exploring the Use of Peer Assessment as a Vehicle for Closing the Gap between Feedback Given and Feedback Used." *Assessment & Evaluation in Higher Education* 35 (5): 551–564.

Eisner, E. W. 1985. *The Art of Educational Evaluation: A Personal View*. London: Falmer Press.

Freire, P. 2000. *Pedagogy of the Oppressed*. New York: Continuum International Publishing Group.

von Glasersfeld, E. 2007. "Learning as a Constructive Activity." In *Key Works in Radical Constructivism*, edited by M. Larochelle, 3–19. Rotterdam: Sense Publishers.

Gourlay, S. 2006a. "Conceptualizing Knowledge Creation: A Critique of Nonaka's Theory." *Journal of Management Studies* 43 (7): 1415–1436.

Gourlay, S. 2006b. "Towards Conceptual Clarity for 'Tacit Knowledge': A Review of Empirical Studies." *Knowledge Management Research and Practice* 4 (1): 60–69.

Hattie, J., and H. Timperley. 2007. "The Power of Feedback." *Review of Educational Research* 77 (1): 81–112.

Kift, S. M., and K. E. Moody. 2009. "Harnessing Assessment and Feedback in the First Year to Support Learning Success, Engagement and Retention." In *ATN Assessment Conference*. Melbourne, Victoria: RMIT University.

Kift, S., K. Nelson, and J. Clarke. 2010. "Transition Pedagogy: A Third Generation Approach to Fye – A Case Study of Policy and Practice for the Higher Education Sector." *The International Journal for the First Year in Higher Education* 1 (1): 1–20.

Klein, J. H. 2008. "Some Directions for Research in Knowledge Sharing." *Knowledge Management Research and Practice* 6 (1): 41–46.

Krause, K.-L., and H. Coates. 2008. "Students' Engagements in First-year University." *Assessment & Evaluation in Higher Education* 33 (5): 493–505.

Lave, J., and E. Wenger. 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.

Li, L., X. Liu, and A. L. Steckelberg. 2010. "Assessor or Assessee: How Student Learning Improves by Giving and Receiving Peer Feedback." *British Journal of Educational Technology* 41 (3): 525–536.

Liu, N.-F., and D. Carless. 2006. "Peer Feedback: The Learning Element of Peer Assessment." *Teaching in Higher Education* 11 (3): 279–290.

MacDonald, K. 2011. "A Reflection on the Introduction of a Peer and Self Assessment Initiative." *Practice and Evidence of Scholarship of Teaching and Learning in Higher Education* 6 (1): 27–42.

Magin, D. 2001. "Reciprocity as a Source of Bias in Multiple Peer Assessment of Group Work." *Studies in Higher Education* 26 (1): 53–63.

McMahon, T. 2010. "Peer Feedback in an Undergraduate Programme: Using Action Research to Overcome Students' Reluctance to Criticise." *Educational Action Research* 18 (2): 273–287.

Mulder, R. A., C. Baik, R. Naylor, and J. Pearce. 2013. "How Does Student Peer Review Influence Perceptions, Engagement and Academic Outcome? A Case Study." *Assessment & Evaluation in Higher Education*. doi:10.1080/02602938.2013.860421.

Mulder, R. A., and J. M. Pearce. 2007. "PRAZE: Innovating Teaching through Online Peer Review." Paper presented at the Ascilite Conference, Nangyang Technological University, Singapore, December 2–5.

Nelson, K. J., J. E. Clarke, S. Kift, and T. A. Creagh. 2011. *Trends in Policies, Programs and Practices in the Australasian First Year Experience Literature 2000–2010. The First Year in Higher Education Research Series on Evidence-based Practice, No. 1*. Brisbane: Queensland University of Technology.

Nicol, D. J., and D. Macfarlane-Dick. 2006. "Formative Assessment and Self-regulated Learning: A Model and Seven Principles of Good Feedback Practice." *Studies in Higher Education* 31 (2): 199–218.

Nicol, D., A. Thomson, and C. Breslin. 2014. "Rethinking Feedback Practices in Higher Education: A Peer Review Perspective." *Assessment & Evaluation in Higher Education* 39 (1): 102–122.

O'Donovan, B., M. Price, and C. Rust. 2004. "Know What I Mean? Enhancing Student Understanding of Assessment Standards and Criteria." *Teaching in Higher Education* 9 (3): 325–335.

O'Donovan, B., M. Price, and C. Rust. 2008. "Developing Student Understanding of Assessment Standards: A Nested Hierachy of Approaches." *Teaching in Higher Education* 13 (2): 205–217.

Orsmond, P., S. Merry, and K. Reiling. 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.

Papinczak, T., L. Young, and M. Groves. 2007. "Peer Assessment in Problem-based Learning: A Qualitative Study." *Advances in Health Sciences Education* 12 (2): 169–186.

Patton, C. 2011. "'Some Kind of Weird, Evil Experiment': Student Perceptions of Peer Assessment." *Assessment & Evaluation in Higher Education* 37 (6): 1–13.

Polanyi, M. 1958. *Personal Knowledge: Towards a Post-critical Philosophy*. London: Routledge and Kegan Paul.

Polanyi, M. 1962. "Tacit Knowing: Its Bearing on Some Problems of Philosophy." *Reviews of Modern Physics* 34 (4): 601–616.

Robinson, R. 2001. "Calibrated Peer Review: An Application to Increase Student Reading and Writing Skills." *The American Biology Teacher* 63 (7): 474–480.

Rubin, R. F., and T. Turner. 2012. "Student Performance on and Attitudes toward Peer Assessments on Advanced Pharmacy Practice Experience Assignments." *Currents in Pharmacy Teaching and Learning* 4 (2): 113–121.

Rust, C. 2001. *A Briefing on Assessment of Large Groups. Learning and Teaching Support Network Generic Centre. Assessment Series No. 12*. York: Learning and Teaching Support Network.

Rust, C., B. O'Donovan, and M. Price. 2005. "A Social Constructivist Assessment Process Model: How the Research Literature Shows us This Could be Best Practice." *Assessment & Evaluation in Higher Education* 30 (3): 231–240.

Rust, C., M. Price, and B. O'Donovan. 2003. "Improving Students' Learning by Developing Their Understanding of Assessment Criteria and Processes." *Assessment & Evaluation in Higher Education* 28 (2): 147–164.

Sadler, D. R. 2010. "Beyond Feedback: Developing Student Capability in Complex Appraisal." *Assessment & Evaluation in Higher Education* 35 (5): 535–550.

Smith, H., A. Cooper, and L. Lancaster. 2002. "Improving the Quality of Undergraduate Peer Assessment: A Case for Student and Staff Development." *Innovations in Education and Teaching International* 39 (1): 71–81.

Topping, K. 1998. "Peer Assessment between Students in Colleges and Universities." *Review of Educational Research* 68 (3): 249–276.

Venables, A., and R. Summit. 2003. "Enhancing Scientific Essay Writing Using Peer Assessment." *Innovations in Education & Teaching International* 40 (3): 281–290.

Vickerman, P. 2009. "Student Perspectives on Formative Peer Assessment: An Attempt to Deepen Learning?" *Assessment & Evaluation in Higher Education* 34 (2): 221–230.

Weber, R. P. 1990. *Basic Content Analysis*. 2nd ed. Newbury Park, CA: Sage.

van Zundert, M., D. Sluijsmans, and J. van Merriënboer. 2010. "Effective Peer Assessment Processes: Research Findings and Future Directions." *Learning and Instruction* 20 (4): 270–279.