

**Peer and Self-Assessment as a Means to Improve Levels of Reflection in Students' Journal Writing**

Journal:	<i>Reflective Practice</i>
Manuscript ID:	CREP-2012-0102
Manuscript Type:	Original Paper
Keywords:	Reflection, Reflective Practice, Reflective Learning, Peer Assessment, Self Assessment

SCHOLARONE™  
Manuscripts

Review Only

## Peer and Self-Assessment as a Means to Improve Levels of Reflection in Students' Journal Writing

Much has been written about ways in which we might help students to improve the level of reflection that can be found in their Journal Writing. Formative peer and self-assessment has been cited as a way for students to develop a clearer understanding of what is required to write reflectively and can alleviate some of the difficulties associated with staff assessing student journals. In order to evaluate the efficacy, reliability and validity of peer and self-assessment, a cohort of architecture students were asked to assess each other's on-line reflective blogs on two occasions during an academic year. These were rated in terms of the level and focus of reflection using a standard assessment rubric. It was found that there was reasonable consistency of ratings between those reviewing the same student's work. There was less consistency between an individual's own assessment and the assessment of their work by their peers. There was also a significant improvement in the level of reflection recorded in the peer-assessment between the first and second review. It appears that this improvement resulted from the process of undertaking the peer assessment, rather than through the utilisation of the feedback it produced.

Keywords: Reflection, Peer-Assessment, Self-Assessment

### Literature Review

Whilst the literature on reflective practice is broad and extensive, definitions of what is meant by reflection vary between authors (McCarthy, 2011; Moon, 1999). Nevertheless, there does seem to be some degree of agreement that reflection can take a number of forms, and manifest itself at a series of cognitive levels, usually ranging from pure description (no reflection) to something that is critical and evaluative, enabling learners to gain new perspectives (Bain, Ballantyne, Packer, & Mills, 1999; Hatton & Smith, 1995; Kember et al., 1999; Mezirow, 1991; Moon, 1999). Given that it is often argued that students who engage in higher levels of reflection are likely to show benefits in terms of their learning, it is somewhat disappointing to note many reports that suggest that much of the supposedly reflective writing by students, is little more than a description. Dymont and O'Connell (2011)

1  
2  
3 reviewed 11 studies of situations where students were asked to write learning journals. They  
4  
5 found that in only two of these cases, were the authors content that students were reaching  
6  
7 high levels of reflection. In five cases, levels of reflection were seen to be minimal. Whilst  
8  
9 Dymont and O'Connell recognised that there was variation in the ways in which reflection  
10  
11 was assessed in each of the separate studies, they were able to conclude with reasons why  
12  
13 some journaling exercises might have been more effective than others. These include  
14  
15 whether journals were mandatory, the level of introductory support given to students, the  
16  
17 length of the journal keeping period, the contribution of the journal to the overall mark, the  
18  
19 nature of the relationship between journal keeper and tutor and the provision of formative  
20  
21 feedback to students whilst undertaking the journaling process. Bean and Stevens (2002) also  
22  
23 commented on apparent low levels of reflection in a number of studies of pre-service teachers  
24  
25 and purported the need to provide 'scaffolding' in the form of prompts, feedback and  
26  
27 dialogue to support students in becoming reflective. Bain et al (2002) in response to similar  
28  
29 concerns, evaluated how feedback might help students develop reflective skills for learning  
30  
31 journals. They distinguished between feedback that discussed the topic being reflected upon  
32  
33 (the focus of reflection) and the level of reflection generated. They concluded that whilst  
34  
35 providing students with formative feedback provided general benefits for the improvement of  
36  
37 their writing, feedback that particularly focused on the level of reflection was particularly  
38  
39 effective. Samuels and Betts (2007) also recognised that reflective writing was often overly  
40  
41 descriptive and that teacher support was needed to help students deepen their reflection.  
42  
43 They suggested that the provision of feedback was a key means to help students achieve this.  
44  
45  
46  
47  
48

49  
50 The benefits of giving formative feedback to students, is generally acknowledged as  
51  
52 an important aspect of student learning (Black & Wiliam, 1998; Yorke, 2003), Effective  
53  
54 feedback helps students clarify what is required of them; enables them to re-conceptualise  
55  
56 and reconsider their understanding, skills and ideas; encourages them to become self-critical  
57  
58  
59  
60

1  
2  
3 and can improve motivation and self-esteem (Nicol & Macfarlane-Dick, 2006).  
4  
5 Nevertheless, providing formative feedback can be time consuming to generate, particularly  
6  
7 where large cohorts are involved. Furthermore, if students anticipate their reflective writing  
8  
9 will be assessed, it can reduce their willingness to write openly and introduce an element of  
10  
11 artificiality into their reflection (Boud & Walker, 1998; Stewart & Richardson, 2000;  
12  
13 Sumsion & Fleet, 1996). Creme (2005) attempts to weigh up the arguments between the  
14  
15 motivating effects of assessment, where students will only take the work seriously if it  
16  
17 counts, and the dangers of assessment constraining the students reflections. She argues that  
18  
19 peer and self-assessment, at formative stages might represent an acceptable compromise,  
20  
21 although recognising that in her observations, students peer feedback lacked guidance on how  
22  
23 to develop the writing further.  
24  
25  
26

27  
28 An important aspect of giving feedback is getting students to become self-critical of  
29  
30 their work (Nicol & Macfarlane-Dick, 2006). A clear, common understanding of assessment  
31  
32 criteria is one way to develop this (Cartney, 2010; Rust, Price, & O'Donovan, 2003) and can  
33  
34 avoid a possible mismatch between the conception of feedback given to students, and the  
35  
36 students conceptions of the feedback they receive (Orsmond & Merry, 2011). Falchikov and  
37  
38 Goldfinch (2000) argue that peer assessment more reliably matches staff ratings when clear,  
39  
40 but simple assessment criteria are provided. Rust et Al. (2003) argue that for this to occur  
41  
42 effectively, students and staff need to develop a shared understanding of the meaning of the  
43  
44 assessment criteria. Sadler (1989) argues that students need to develop the same evaluative  
45  
46 skills as their tutors, so that they can conceptualise the feedback from a similar perspective.  
47  
48

49  
50 Ellis (2001) evaluates the benefits of introducing peer assessment into the process of  
51  
52 writing a practice-based reflective journal, and concludes that this process helps students who  
53  
54 are not naturally reflective develop the skills necessary for their own writing. Samuels and  
55  
56 Betts(2007) also note the benefits of self-reflection in helping to develop student's journal  
57  
58  
59  
60

1  
2  
3 writing, suggesting a series of questions or prompts that can be used to help them 'reflect on  
4 their reflection'. Dunning et al (2004) however, caution against self-assessment, observing  
5 that often judgments show elements of overconfidence on the part of the assessor. They  
6 argue that peer assessment remains a more reliable means for assessing work.  
7  
8  
9  
10

### 11 *The Study*

12  
13 This paper evaluates the introduction of peer and self-assessment to the process of journal  
14 writing in order to help students develop their reflective abilities. Its subjects are a cohort of  
15 64 Architecture students who are undertaking a placement year in professional practice; in  
16 many cases their first experience of professional practice. Whilst the research focusses  
17 wholly on architecture students, its findings should be transferable to other professional areas.  
18  
19  
20  
21  
22  
23  
24  
25  
26

27 As a course requirement, students are expected to undertake 9 months work in an  
28 architectural (or related) practice. During their time, they are required to produce a reflective  
29 e-portfolio related to their experience in practice. This e-portfolio is assessed, and constitutes  
30 one third of the marks for the entire year. To maintain fairness, the criteria for assessment is  
31 based entirely on the level of reflection rather than the subject of the student's writing, which  
32 can vary depending on the student's employment circumstances.  
33  
34  
35  
36  
37  
38  
39

40 The principal assessed component to the e-portfolio is a reflective blog that the  
41 students are expected to write. Students are provided with guidance which suggests they  
42 consider what they regarded as challenging and why; their feelings, beliefs and assumptions  
43 and how these might have changed; the relationship between their experience and their initial  
44 expectations of practice; their roles and relationships to those of their colleagues; how they  
45 addressed the challenges encountered; the support and feedback they had and needed and  
46 what could be done better next time. These prompts are designed to steer the students away  
47 from purely descriptive writing and to encourage something that is more reflective.  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 In previous years, the author has been responsible for reading all of the student's e-  
4 portfolios on four occasions during the academic year. Given that a typical cohort size is  
5 approximately 60 students, this proved an onerous task, and often led to delays in feedback  
6 being returned to the students. Furthermore, reliability and consistency of comments could  
7 not be guaranteed. Although end of year course evaluations always suggested that students  
8 value receiving feedback on their work, experience also showed that there appeared to be  
9 very little improvement in the quality of the students' work following the formative feedback.  
10  
11

12  
13  
14  
15  
16  
17  
18  
19 Recognising concerns about consistency in terms of comments and marks, the author  
20 conducted a detailed analysis of the reflective writing of a previous cohort [Reference  
21 Removed]. This led to the generation of the rubric that can be used to classify the students'  
22 reflection (Appendix 1). As with the taxonomies developed by Bain et al. (1999) this  
23 distinguishes between the level of reflection provided by the student, and the particular focus  
24 of reflection – in this case between focussing on their own personal development, the specific  
25 projects they are working on in their offices, or the environment, practice or broader  
26 profession in which they are currently working. The rubric provides generic descriptors for  
27 the differing levels of reflection within each focus. This can be used by both students and  
28 teachers as a point of reference to measure level and focus of reflection. It was hoped that by  
29 using it to carry out peer and self assessment, it would enable the students to better  
30 understand what was meant by reflection.  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44

45  
46 The students were issued with copies of the assessment rubric and a set of explanatory  
47 notes. They were allocated two other students whose work they had to review using the  
48 rubric in addition to reviewing their own work. This was done on two occasions – the first  
49 four months after starting the programme, and the second three months after that. To enable  
50 the students to be exposed to a fuller range of their colleagues' work, they reviewed different  
51 students on each occasion. The students were asked to read their allocated reflections a  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 paragraph at a time, and allocate each paragraph to one (or more) of the nine categories in the  
4  
5 rubric. The nine categories were a combination of three levels of reflection (high, medium  
6  
7 and low), and three possible foci of reflection (project, person and practice). In some cases  
8  
9 paragraphs may feature text that falls into more than one category, and students were told that  
10  
11 they could allocate a paragraph to more than once. Students were asked to count the number  
12  
13 of paragraphs in each category and report these back. In a similar study, Bell et al (2010)  
14  
15 found that coding individual paragraphs gave a better impression of overall levels of  
16  
17 reflection than making a holistic judgement across individual pieces of writing. They found  
18  
19 that students predominately wrote the majority of their work at low levels of reflection, but  
20  
21 most wrote a small number of paragraphs at a higher level (some very high). They argued  
22  
23 that by taking a holistic view, reviewers might lose sight of this high level of reflection.  
24  
25 Similarly Falchikov and Goldfinch (2000) through a meta-analysis of studies comparing tutor  
26  
27 marks with peer assessment marks, found less consistency where a global judgement had to  
28  
29 be made compared to those where explicit criteria were provided with several dimensions.  
30  
31 Therefore, the decision was made to avoid asking students to make a global judgement, but to  
32  
33 rely on them categorising individual paragraphs.  
34  
35  
36  
37

38  
39 In addition to the assigning of paragraphs to categories the students were asked to rate  
40  
41 each portfolio in terms of how thorough it was, how well presented it was and how well  
42  
43 written it was. These were all rated using a scale of 1-5. Finally the students were asked to  
44  
45 write some optional free text advice for the students. The students then entered the total  
46  
47 number of paragraphs counted in each category into an on-line Google form.  
48

49  
50 The peer and self-review results were presented back to the students visually using a  
51  
52 graph (Figure 1). The graph was designed to quickly show students how their portfolios  
53  
54 compared to the cohort as a whole and how their own rating and peer ratings compared. The  
55  
56 graph used an algorithm to create a single score for the underlying level of reflection for each  
57  
58  
59  
60

1  
2  
3 of the three foci within the students writing from the overall paragraph counts. A set of target  
4  
5 scores was also created to indicate to the students how we would wish the students to develop  
6  
7 their portfolios. The target score was the same for each student but was a somewhat arbitrary  
8  
9 means to focus their writing. It placed greater emphasis on encouraging the students to write  
10  
11 more in the practice and personal areas, rather than to describe their projects. The students  
12  
13 could also see their peers comments on their work. The tutor also reviewed each of the  
14  
15 students work and provided comments on the extent to which the peer reviews seemed a fair  
16  
17 reflection of the work (generally it was felt that they were).  
18  
19

20  
21 [Figure 1 here]  
22  
23

## 24 25 **Findings**

26  
27 As mentioned previously, the students reviewed each other's work on two occasions. As the  
28  
29 students are based out in architectural practice, rather than in the university, some found it  
30  
31 difficult to prioritise academic work over their professional work. As a result, at the time of  
32  
33 the first review, a number had barely started completing their blogs, and so it was impossible  
34  
35 for other students to review them. All students were still expected to review two other  
36  
37 students, so for the first review there could be up to 4 students reviewing a peer's work. By  
38  
39 the second review, there were generally only two reviewers as most students had written  
40  
41 something. At the first review 39 students had work to review and at the second review,  
42  
43 there were 64.  
44  
45  
46  
47

48 The data submitted by each student took the form of a count of the number of  
49  
50 paragraphs falling into each of the nine categories in the assessment rubric. It was possible to  
51  
52 use this to measure consistency of rating, and to see if student writing had improved as a  
53  
54 result of the exercise. It should be noted at this point that students were not expected to  
55  
56  
57  
58  
59  
60



1  
2  
3 submit marked up copies of the materials they had assessed so there was no indication of  
4  
5 which paragraphs were marked at each level.  
6  
7

### 8 9 *Consistency of ratings between students*

10  
11 The ratings from each of the reviewers for an individual student were compared to measure  
12 consistency. The number of paragraphs counted by each reviewer was used to calculate the  
13 mean for that student in each of the categories. Consistency was measured by looking at each  
14 reviewer's average deviation from that mean. In such a way an average deviation of zero  
15 suggests that all reviewers allocated exactly the same number of paragraphs to that particular  
16 category for that particular student, an average deviation of less than 1, suggests that the  
17 differences between reviewers was less than one paragraph either side of the mean.  
18  
19

20  
21 The results from both the first and second review showed that between half and two  
22 thirds of reviewers varied in their categorisation by one paragraph or less from the mean for  
23 the student they reviewed (Tables 1 and 2). Approximately 90% of reviewers varied in their  
24 categorisation by two paragraphs or less. Given that the average number of paragraphs  
25 counted was 19 in the first review and 23 in the second review, one or two paragraphs  
26 deviation would appear consistent enough to provide feedback guidance to the students.  
27  
28

29  
30 [Tables 1 and 2 here]  
31  
32

33  
34 The tables suggest that there was some variation in the consistency between the different  
35 classification categories. This might suggest that certain statements in the rubric may have  
36 been more difficult to interpret than others. Nevertheless, these differences were not  
37 replicated between the two reviews suggesting that it is not possible to draw conclusions on  
38 the clarity of the rubric from this data.  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

### *Consistency between peer and self ratings*

The mean number of paragraphs counted in each category for each student during the peer review was compared to the counts, for the students' self-evaluations. At the point of the first review the sample size for this is reduced as some of the students believed (wrongly) that they did not need to record their own self-assessment on the on-line system. 25 students were reviewed by themselves as well as their peers. The percentage of students whose ratings were either identical to their peers, or differed by a number of paragraphs are shown in tables 3 and 4.

[Tables 3 and 4 here]

The self-reviews on the whole appeared to show a greater spread in terms of their deviation from the peer review mean, often with over 20% of the students by more than 4 paragraphs in particular categories. By the second review a greater percentage of students were giving identical ratings to their peer review mean, but at the same time there were some substantial deviations. There was however no clear evidence to suggest that students were consistently under, or over rating themselves with roughly similar numbers of students rating each side of the peer-review mean. Closer inspection of the data suggested that a small number of students provided paragraph counts in their self-reviews that consistently varied from their peer reviewers across all nine categories, but the majority tended to be very close to their peer reviewers (by two paragraphs or less) in all but one or two categories. Those students whose self-ratings varied significantly from their peer ratings in the second review tended to be the same students who had a significant difference in the first review, or those who did not self-review themselves in the first review, suggesting that they hadn't established a mental benchmark of what the standards in the criteria meant. This suggests that some students may have had a different perception as to the level that a paragraph was written at. It may also

1  
2  
3 suggest that the student's intention was to write critically, but the written work was not  
4  
5 interpreted as so by others, or that students were unaware that they were being reflective.  
6  
7

### 8 9 *Impact on Students*

10  
11 For those students whose portfolios were peer-reviewed on both occasions, it was possible to  
12  
13 compare their average ratings at the first assessment and at the second assessment. One  
14  
15 would expect that if the process of peer reviewing had had a positive effect on the writing,  
16  
17 then the students would have a larger proportion of paragraphs categorised at a higher level  
18  
19 of reflection, and a smaller proportion categorised at the lower levels. The data summarised  
20  
21 in Table 5 suggests that this was the case across all three of the foci with effect sizes of  
22  
23 between 0.5 and 0.7. A repeated measures t-test suggested that these changes were  
24  
25 statistically significant to at least the 95% level of certainty. As the total numbers of  
26  
27 paragraphs categorised would vary between the first and second review, a figure for the  
28  
29 proportion of paragraphs in each category was calculated, by dividing the number of  
30  
31 paragraphs in each category, with the total number of paragraphs counted for each student.  
32  
33  
34  
35

36  
37 [Table 5 here]  
38

39  
40 Those students who didn't complete a portfolio in time for the first review were obviously  
41  
42 excluded from this initial analysis. Given that this group still reviewed other people's  
43  
44 portfolios, but did not have their portfolios reviewed on the first attempt, it is useful to  
45  
46 determine whether the group who had been reviewed twice, were writing with higher levels  
47  
48 of reflection following the first review, compared to the others. This was supported  
49  
50 statistically with a significantly higher proportion of paragraphs being written in the high-  
51  
52 project and high-person categories by those had been reviewed twice. One might also expect  
53  
54 that the number of paragraphs coded at low levels would be lower for those whose work had  
55  
56 been reviewed twice (Table 6). Whilst the raw figures suggest this to be the case, the  
57  
58  
59  
60

1  
2  
3 differences were not sufficient to be statistically significant. This slightly mixed result might  
4  
5 be explained by the fact that some students who were not reviewed at the first stage, were still  
6  
7 benefiting from carrying out the peer review process.  
8  
9

10 [Table 6 here]  
11

12 This explanation was supported by additional data collected at the time of the second review.  
13

14 Students were asked to rate on a 5 point likert scale from strongly agree to disagree whether  
15  
16 “Reviewing other students portfolios has helped me to develop my own portfolio” and “The  
17  
18 feedback provided following the first review has helped me to develop my own portfolio”.  
19  
20

21 Approximately 60% of the students either agreed or strongly agreed that the reviewing other  
22  
23 students portfolios had been of benefit, with the majority of the remainder were neutral.  
24  
25

26 However, only 15% of the students agreed or strongly agreed that the feedback provided had  
27  
28 been helpful. Students were also asked to provide free-text comments and these comments  
29  
30 supported the numerical ratings. A number of students commented on how useful the process  
31  
32 of reviewing other students work as and how it helped them to contextualise their experience  
33  
34 against that of other students. It was noted by one student that the process had changed his  
35  
36 approach to writing his own blog, but notes, as suggested above, that the process of reviewing  
37  
38 was more valuable than the feedback provided.  
39  
40  
41  
42

43 “I found reading the blogs of others interesting and enjoyable and it certainly changed  
44  
45 my approach to my own. I think that, even if I had useful feedback, the reviewing  
46  
47 process would remain more useful for the development of my own work.”  
48

## 49 **Conclusions**

50 The analysis above suggests that the process of asking students to peer-review each other’s  
51  
52 work was a useful exercise. The results suggest that there were clear improvements in the  
53  
54 level of reflection provided in the writing of the students following the first stage review.  
55  
56  
57  
58  
59  
60

1  
2  
3 What was also apparent was that it appeared that the process of carrying out the peer and self-  
4  
5 reviews was more valuable than the feedback that the process generated. This concurs with  
6  
7 research by Rust et al (2003) suggesting that a fuller understanding of assessment criteria can  
8  
9 lead to an improvement of student work. Further work by Bell et al (2012) suggests that this  
10  
11 leads to questions as to how students use the assessment criteria. In some cases students use  
12  
13 assessment criteria as a 'recipe' for further improvement; many however use the criteria as a  
14  
15 guide or framework for completing an assessment task. Further research is needed related to  
16  
17 the provided rubric to get a better understanding of how students are using the criteria  
18  
19 provided here. One concern in the present research is the apparent inconsistency of students'  
20  
21 assessment of their own work, by comparison with the consistency of assessment between  
22  
23 peer reviewers. Dunning(2004) argued that self-assessments typically over inflate an  
24  
25 individual's performance, but in this case the results were rather more mixed – with some  
26  
27 students significantly over-estimating themselves, and others under-estimating themselves.  
28  
29 These differences again may be a result of the way that students understand and interpret the  
30  
31 assessment criteria. This may also be an aspect of when students assess their own work; they  
32  
33 supplement what is written with tacit knowledge from their own experiences. They base their  
34  
35 assessment on that entire picture, rather than a more limited view that their peer reviewers  
36  
37 will see.  
38  
39  
40  
41  
42  
43

#### 44 *Limitations and opportunities for further research*

45  
46  
47 With the exception of the brief questionnaire that was completed at the time of the second  
48  
49 submission, the data for this research was largely based on data collected as part of an  
50  
51 assessment exercise. Ethical reasons meant that it was not possible to set up control groups  
52  
53 of students who did not undertake the peer-review exercise, and the comparison groups were  
54  
55 therefore self-selected, based on whether work had been submitted on time. The data  
56  
57  
58  
59  
60

1  
2  
3 collection method was designed to balance simplicity of judgement (by counting paragraphs)  
4  
5 with collection of a rigorous data set that that could be used for evaluation of the exercise and  
6  
7 perhaps this precluded the possibilities of making a more holistic judgement.  
8

9  
10 The research was also limited in that only student ratings were included. Whilst the  
11 author as course tutor was required to grade the work as part of a final assessment  
12 requirement, there was no independent tutor rating that could be compared with the students'  
13 ratings. Furthermore whilst anecdotally, the student work from the cohort studied here did  
14 appear to show higher levels of reflection than in previous year (where no formal peer  
15 assessment took place), comparative data simply did not exist that could confirm this. One  
16 further limitation, already hinted upon is that the student blogs were marked by a member of  
17 staff, and were also made available to the entire cohort to read. These factors may well have  
18 limited the students' willingness to write expressively and openly; perhaps limiting their level  
19 of reflection. This is a difficult issue to address, given that if not assessed, or open, then it is  
20 unlikely that all students would engage in the exercise of writing reflectively (Roberts, 2009).  
21  
22

23  
24 Despite these limitations, the exercise formed a useful pedagogic intervention that has  
25 seemingly improved student outputs. The prepared rubric seemed to be key to this success, by  
26 helping students think about what they should write about, and at what level. Whilst the  
27 focus of this report was a cohort of architecture students its findings are no doubt relevant to  
28 other professional disciplines who are attempting to improve students' level of reflection.  
29  
30  
31  
32  
33

### 34 35 36 37 38 39 40 41 42 43 44 45 46 47 **References**

- 48  
49 Bain, J. D., Ballantyne, R., Packer, J., & Mills, C. (1999). Using Journal Writing to Enhance Student  
50 Teachers' Reflectivity During Field Experience Placements. *Teachers and Teaching*, 5(1), 51 -  
51 73.  
52 Bain, J. D., Mills, C., Ballantyne, R., & Packer, J. (2002). Developing Reflection on Practice Through  
53 Journal Writing: impacts of variations in the focus and level of feedback. *Teachers and*  
54 *Teaching*, 8(2), 171 - 196.  
55 Bean, T. W., & Stevens, L. P. (2002). Scaffolding Reflection for Preservice and Inservice Teachers.  
56 *Reflective Practice*, 3(2), 205-218. doi: 10.1080/14623940220142343  
57  
58  
59  
60

- 1  
2  
3 Bell, A., Kelton, J., McDonagh, N., Mladenovic, R., & Morrison, K. (2010). A critical evaluation of the  
4 usefulness of a coding scheme to categorise levels of reflective thinking. *Assessment &*  
5 *Evaluation in Higher Education*, 36(7), 797-815. doi: 10.1080/02602938.2010.488795  
6 Bell, A., Mladenovic, R., & Price, M. (2012). Students' perceptions of the usefulness of marking  
7 guides, grade descriptors and annotated exemplars. *Assessment & Evaluation in Higher*  
8 *Education*, 1-20.  
9  
10 Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education:*  
11 *Principles, Policy & Practice*, 5(1), 7-74. doi: 10.1080/0969595980050102  
12 Boud, D., & Walker, D. (1998). Promoting reflection in professional courses: The challenge of  
13 context. *Studies in Higher Education*, 23(2), 191-206. doi: 10.1080/03075079812331380384  
14 Cartney, P. (2010). Exploring the use of peer assessment as a vehicle for closing the gap between  
15 feedback given and feedback used. *Assessment & Evaluation in Higher Education*, 35(5), 551-  
16 564. doi: Pii 926292288  
17 10.1080/02602931003632381  
18 Creme, P. (2005). Should student learning journals be assessed? *Assessment & Evaluation in Higher*  
19 *Education*, 30(3), 287-296. doi: 10.1080/02602930500063850  
20 Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed Self-Assessment. *Psychological Science in the*  
21 *Public Interest*, 5(3), 69-106. doi: 10.1111/j.1529-1006.2004.00018.x  
22 Dymont, J. E., & O'Connell, T. S. (2011). Assessing the quality of reflection in student journals: a  
23 review of the research. *Teaching in Higher Education*, 16(1), 81-97. doi:  
24 10.1080/13562517.2010.507308  
25  
26 Ellis, G. (2001). Looking at Ourselves—Self-Assessment and Peer Assessment: Practice examples  
27 from New Zealand. *Reflective Practice*, 2(3), 289-302. doi: 10.1080/1462394012010303  
28 Falchikov, N., & Goldfinch, J. (2000). Student Peer Assessment in Higher Education: A Meta-Analysis  
29 Comparing Peer and Teacher Marks. *Review of Educational Research*, 70(3), 287-322. doi:  
30 10.3102/00346543070003287  
31 Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and  
32 implementation. *Teaching and Teacher Education*, 11(1), 33-49.  
33 Kember, D., Jones, A., Loke, A., Mckay, J., Sinclair, K., Tse, H., . . . Yeung, E. (1999). Determining the  
34 level of reflective thinking from students' written journals using a coding scheme based on  
35 the work of Mezirow. *International Journal of Lifelong Education*, 18(1), 18 - 30.  
36 McCarthy, J. (2011). Reflective Writing, Higher Education and Professional Practice. *Journal of*  
37 *Education in the Built Environment*, 6(1), 15.  
38 Mezirow, J. (1991). *Transformative dimensions of adult learning* (1st ed ed.). San Francisco: Jossey-  
39 Bass.  
40 Moon, J. A. (1999). *Reflection in learning and professional development : theory and practice*.  
41 London: Kogan Page.  
42 Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model  
43 and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218.  
44 doi: 10.1080/03075070600572090  
45 Orsmond, P., & Merry, S. (2011). Feedback alignment: effective and ineffective links between tutors'  
46 and students' understanding of coursework feedback. *Assessment & Evaluation in Higher*  
47 *Education*, 36(2), 125-136. doi: 10.1080/02602930903201651  
48 Roberts, A. (2009). Encouraging reflective practice in periods of professional workplace experience:  
49 the development of a conceptual model. *Reflective Practice*, 10(5).  
50 Rust, C., Price, M., & O'Donovan, B. (2003). Improving Students' Learning by Developing their  
51 Understanding of Assessment Criteria and Processes. *Assessment & Evaluation in Higher*  
52 *Education*, 28(2), 147-164. doi: 10.1080/02602930301671  
53 Sadler, D. R. (1989). FORMATIVE ASSESSMENT AND THE DESIGN OF INSTRUCTIONAL-SYSTEMS.  
54 *Instructional Science*, 18(2), 119-144. doi: 10.1007/bf00117714  
55  
56  
57  
58  
59  
60

- 1  
2  
3 Samuels, M., & Betts, J. (2007). Crossing the threshold from description to deconstruction and  
4 reconstruction: using self-assessment to deepen reflection. *Reflective Practice, 8*(2), 269-  
5 283. doi: 10.1080/14623940701289410  
6  
7 Stewart, S., & Richardson, B. (2000). Reflection and its Place in the Curriculum on an Undergraduate  
8 Course: Should it be assessed? *Assessment & Evaluation in Higher Education, 25*(4), 369-380.  
9 doi: 10.1080/713611443  
10  
11 Sumsion, J., & Fleet, A. (1996). Reflection: can we assess it? Should we assess it? *Assessment &*  
12 *Evaluation in Higher Education, 21*(2), 121-130. doi: 10.1080/0260293960210202  
13  
14 Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the  
15 enhancement of pedagogic practice. *Higher Education, 45*(4), 477-501. doi:  
16 10.1023/a:1023967026413  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



**Appendix 1: Summary of the nature of students' writing in the nine categories of level and focus of reflection. A rubric for assessment**

	<b>High Reflection</b>	<b>Medium Reflection</b>	<b>Low Reflection</b>
<b>Project Focus</b>	<p><i>Critical Review</i></p> <p>Provides a critical review of the building or project.</p> <p><i>Role Evaluation</i></p> <p>Students evaluate their contribution towards the project.</p>	<p><i>Explanation</i></p> <p>Provides information about the project, but makes links to the reasons why things are as they are.</p> <p>Outlines what they have found challenging and why.</p> <p>Outlines how they contributed towards the project and interacted with others.</p>	<p><i>Description</i></p> <p>Describes aspects of project, for instance client requirements, schedules etc.</p> <p>Outlines what they have done.</p> <p>Outlines how things were done on the project (not necessarily what they did).</p>
<b>Profession or Practice Focus</b>	<p><i>Theorisation</i></p> <p>Proposes some novel theory about how practices work.</p> <p>Critically evaluates process.</p> <p>Improves process. Expresses an opinion on success or failure, and gives reasoning.</p>	<p><i>Generalisation</i></p> <p>Generalises observations within their practice to wider practice. Makes comparison with other prior experience.</p>	<p><i>Observation</i></p> <p>Observes nature of individual's practice/office. Describes context.</p> <p>Describes process observed within practice.</p>
<b>Personal Development focus</b>	<p><i>Transformation</i></p> <p>Highlights new insights.</p> <p>Evaluates success in meeting challenges.</p> <p>Sees their experience in a new way.</p> <p>Overall evaluation of impact of placement, very much generalised.</p> <p>Shows high degree of introspection in terms of understanding how they learn.</p>	<p><i>Self Evaluation</i></p> <p>Reflects on how they have developed.</p> <p>Shows emotions – but explains why these occur.</p> <p>Explains how they have improved. Extracts skills rather than broad aspects.</p> <p>Links practice and theory.</p> <p>Evaluates benefit of experience and highlights carefully areas for improvement. What would they do better next time? Generally project/task specific.</p> <p>Outlines what has helped them to learn.</p>	<p><i>Response</i></p> <p>Expresses a simple emotion (like, dislike). Describes difficulty encountered, a challenging task.</p> <p>Lists what has been learned or needs to have been learned.</p> <p>Lists things they need to do based on standard pre-conception, range of experience within plan of work etc.</p>

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For Peer Review Only

Table 1: consistency between reviewers in first review

		Focus and Level								
		Project			Person			Practice		
		High	Med	Low	High	Med	Low	High	Med	Low
Average deviation (paragraphs)	0	26%	8%	10%	18%	13%	8%	18%	3%	15%
	0-1	36%	59%	41%	59%	49%	49%	49%	56%	41%
	1-2	28%	23%	26%	15%	26%	26%	28%	31%	38%
	2-3	5%	8%	10%	0%	8%	8%	5%	8%	3%
	3-4	5%	3%	5%	5%	3%	8%	0%	0%	0%
	>4	0%	0%	8%	3%	3%	3%	0%	3%	3%

Note: percentages represent the proportion of reviewers deviating from the mean number of paragraphs counted for each student reviewed.

Table 2: consistency between reviewers in second review

		Focus and Level								
		Project			Person			Practice		
		High	Med	Low	High	Med	Low	High	Med	Low
Average deviation (paragraphs)	0	17%	11%	22%	11%	11%	14%	14%	25%	17%
	0-1	50%	42%	31%	50%	50%	50%	53%	42%	50%
	1-2	19%	36%	28%	25%	33%	22%	22%	25%	19%
	2-3	8%	8%	8%	11%	3%	6%	3%	6%	6%
	3-4	0%	3%	6%	0%	3%	3%	3%	3%	3%
	>4	6%	0%	6%	3%	0%	6%	6%	0%	6%

Note: percentages represent the proportion of reviewers deviating from the mean number of paragraphs counted for each student reviewed.

Table 3: consistency between self and peer reviewers in first review

		Focus and Level								
		Project			Person			Practice		
		High	Med	Low	High	Med	Low	High	Med	Low
Absolute difference from peer review mean (paragraphs)	0	12%	12%	12%	12%	8%	12%	16%	12%	8%
	0-1	64%	32%	20%	40%	44%	44%	52%	20%	52%
	1-2	16%	20%	20%	36%	24%	12%	12%	28%	20%
	2-3	4%	16%	8%	4%	16%	16%	12%	24%	12%
	3-4	4%	0%	12%	0%	4%	8%	0%	16%	0%
	>4	0%	20%	28%	8%	4%	8%	8%	0%	8%

Note: percentages represent the proportion of self- reviews that differed from the mean paragraph count for the peer review of each student reviewed.

Table 4: consistency between self and peer reviewers in second review

		Focus and Level								
		Project			Person			Practice		
		High	Med	Low	High	Med	Low	High	Med	Low
Absolute difference from peer review mean (paragraphs)	0	20%	13%	13%	23%	3%	15%	13%	13%	20%
	0-1	30%	25%	20%	28%	30%	28%	30%	30%	35%
	1-2	18%	23%	28%	10%	23%	23%	25%	23%	20%
	2-3	10%	5%	10%	13%	13%	13%	10%	20%	10%
	3-4	3%	3%	5%	5%	10%	5%	5%	5%	0%
	>4	20%	33%	25%	23%	23%	18%	18%	10%	15%

Note: percentages represent the proportion of self- reviews that differed from the mean paragraph count for the peer review of each student reviewed.

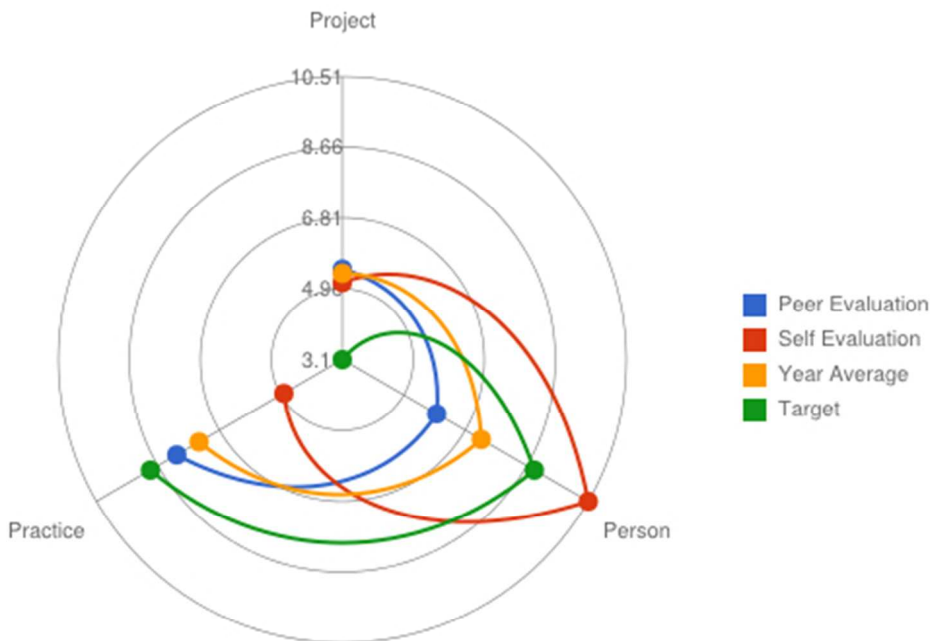
Table 5: proportion of paragraphs written at each level and focus between the first and second review (N=35).

	Level and Focus								
	Project			Person			Practice		
	High	Med	Low	High	Med	Low	High	Med	Low
<b>1st review</b>	7%	13%	15%	7%	15%	11%	7%	13%	11%
<b>2nd Review</b>	10%	15%	10%	11%	15%	8%	11%	13%	7%
<b>t value</b>	2.001	1.17	-2.26	2.897	-0.2	-3.27	1.834	-0.23	-1.94
<b>p (1 tail)</b>	0.027	0.125	0.015	0.003	0.422	0.001	0.038	0.409	0.03
<b>Effect size (cohen's d)</b>	0.489	0.244	-0.57	0.616	-0.05	-0.69	0.448	-0.05	-0.53

Table 6: proportion of paragraphs written at each level and focus in the second review for those who had received feedback after the first review and those who had not

	Level and Focus								
	Project			Person			Practice		
	High	Med	Low	High	Med	Low	High	Med	Low
<b>reviewed at stage 1 (N=35)</b>	10%	15%	10%	11%	15%	8%	11%	13%	7%
<b>not reviewed at stage 1 (N=19)</b>	6%	17%	12%	8%	18%	7%	8%	13%	10%
<b>t value</b>	2.51	-1.09	-0.86	1.98	-1.72	0.30	1.05	-0.01	-1.23
<b>p (1 tail)</b>	0.01	0.14	0.20	0.03	0.05	0.38	0.15	0.49	0.11
<b>Effect size (cohen's d)</b>	0.642	-0.36	-0.27	0.525	-0.53	0.098	0.319	-0	-0.36

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



example chart as returned to students  
136x98mm (96 x 96 DPI)

view Only