



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Gan, MJS;Hattie, J

Title:

Prompting secondary students' use of criteria, feedback specificity and feedback levels during an investigative task

Date:

2014-10-14

Citation:

Gan, M. J. S. & Hattie, J. (2014). Prompting secondary students' use of criteria, feedback specificity and feedback levels during an investigative task. *Instructional Science*, 42 (6), pp.861-878. <https://doi.org/10.1007/s11251-014-9319-4>.

Persistent Link:

<https://hdl.handle.net/11343/282779>

Running head: Prompting peer feedback

Prompting secondary students' Use of Criteria, Feedback Specificity and  
Feedback Levels during an Investigative task

### **Abstract**

This study investigates the effects of prompting on secondary students' written peer feedback in chemistry investigation reports. In particular, we examined students' feedback features in relation to the use of criteria, feedback specificity, and feedback levels. A quasi-experimental pre-test post-test design was adopted. Reviewers in the prompted condition were provided with question prompts that asked them to pose written feedback to their peers on what they did or did not do well and suggestions for improvement, while reviewers in the unprompted condition gave written peer feedback without prompts. The findings showed that prompted peer feedback has a significant effect on the number of comments related to Knowledge of errors, Suggestions for improvement and Process level feedback. This study supports the view that prompting peer feedback in the use of criteria, feedback specificity and feedback levels opens up opportunity for reviewers to engage more meaningfully with peer feedback in report writing tasks.

**Keywords** Peer feedback; Prompts; Feedback specificity; Feedback levels; Science investigation

## Prompting Reviewers' Use of Criteria, Feedback Specificity and Feedback Levels during an Investigative task

### Introduction

A growing line of feedback research suggests that peers play a crucial role in changing the way students learn in the classroom (Black, Harrison, Lee, Marshall, & Wiliam, 2003; Gibbs & Simpson, 2004; Topping, 2005). Peer feedback has been recognised as a key element in supporting this interaction process, which usually takes the form of reciprocal peer reviews and involves 'a communication process through which learners enter into dialogues related to performance and standards' (Ngar-Fun & Carless, 2006, p. 280). Involving students in peer assessment and peer review enables students to take an active role in managing their own learning (Andrade, 2010), helps in enhancing students' self-assessment skills (Boud, 1995; Boud, Cohen, & Sampson, 1999) and has the potential to improve learning of subject matter (Falchikov, 2001, 2005).

Despite its potential benefits to students, researchers and educators are quick to point out that not all students provide such elaborations or quality feedback (Lockhart & Ng, 1995; Strijbos, Narciss, & Dünnebier, 2010). Often the more able, the more committed, and the more verbal students provide greater elaboration and critical feedback and thus, are more advantaged in peer interactions. There can be resistance to using students to provide feedback and being involved in peer assessment; for example, concerns about the reliability of students' comments, grading or marking, power relations among peers and with teachers; the fact that some students can fail to participate (social loafing), freeloading off others, be impacted by friendship bonds, power relations, or collusion (e.g. Kim, 2009; Ngar-Fun & Carless, 2006). To date, most of the research findings have been derived from studies in higher education, and empirical studies on understanding students' peer feedback skills within the secondary school context are scarce (e.g. Gielen, Peeters, Dochy, Onghena & Struyven, 2010).

Given this complexity, our study investigated how the use of prompts can better support reviewers' in posing feedback that includes features to help their peers in report writing. In particular, the purpose of this study was to examine secondary students' prompted and unprompted peer feedback in relation to their use of criteria, feedback specificity and feedback levels for report writing improvement.

### Peer feedback features

Past studies have emphasised the need to support peer feedback in the classrooms (Cho & MacArthur, 2010; Lundstrom & Baker, 2009; Prins, Sluijsmans, & Kirschner, 2006; Rollinson, 2005). Teachers planning to use peer review are encouraged to deliberately teach students feedback skills, structure classrooms to share this expertise, and make specific interventions to ensure all students can benefit from these peer interactions. More importantly, instructional support and training that target specific feedback features appear to have a more positive effect on students' understanding and use of feedback (Gielen et al., 2010; Hattie & Timperley, 2007; Min, 2005; Nelson and Schunn, 2009; Sluijsmans, Brand-Gruwel, van Merriënboer, and Bastiaens, 2002; Van Steendam, Rijlaarsdam, Sercu, & Van den Bergh, 2010). Drawing from these studies and other peer feedback literature, we identified and proposed the use of criteria, specificity and levels of feedback as key features for studying peer feedback support in the classroom.

**Use of Criteria.** In peer assessment literature, there is a particular focus on the negotiation about performance criteria (Falchikov & Goldfinch, 2000; Orsmond, Merry, & Reiling, 2000, 2002; Topping, 1998) which is seen as important and as a pre-requisite for both the giver and the receiver of peer feedback. A key argument here is that engaging students in using criteria for peer review would help them to understand what it is that constitutes high quality work (Althausser & Darnall, 2001; Frederiksen & White, 1997; Ploegh, Tillema, & Segers, 2009). For example, Sluijsmans and Van Merrienboer (2000) conducted a literature review and expert interviews on the peer assessment skills that resulted in a peer assessment model in which three main skills are taken into account. These skills are: (1) defining assessment criteria—thinking about what is required and referring to the product or process; (2) judging the performance of a peer reflecting upon and identifying the strengths and weaknesses in a peer's product and writing an assessment report; and (3) providing feedback for future learning—giving constructive feedback about the work produced by a peer. Sluijsmans et al. (2002) incorporated these three skills in training student teachers on peer assessment tasks and demonstrated that explicit training on defining, developing, and using criteria in peer assessment resulted in more constructive comments, higher structure and fewer naive words used.

Involving students in using criteria is not without its challenges. Often, what counts as 'evidence' may vary depending on teachers' expectations and understandings; the unfamiliar

context and representation may also mean that students struggled to decode the feedback message (Sadler, 2009). Studies on expert revisers versus novice writers continued to reveal that less competent reviewers lacked experience with revision, have limited knowledge of revision criteria and failed to use effective revision strategies such as critical reading, detecting and diagnosing problems and other procedural knowledge needed in the revision process (Cho & Schunn, 2007). Thus, explicit instruction with scaffolding tools such as peer review sheets, prompting questions and graphic organisers that target evaluative criteria becomes a necessary part of the peer review process (van Steendam et al., 2010). For example, a detailed case study of a student giving and receiving peer feedback found that it was through a continuous process of reviewing a range of other peers' work that shaped and developed her understanding of quality in written work (McConlogue, 2014). By composing peer feedback and articulating judgements, this student was able to progress towards a higher quality work which would otherwise be seen as beyond her own abilities if she relied on an excellent 'model answer' (p. 10).

**Feedback specificity.** When supporting the specificity of peer feedback, there are some commonalities in the indicators used. Peer feedback is often examined in terms of identifying a "learning gap" between current and desired performance (Ramaprasad, 1983; Sadler, 1989) and how best to use the feedback information to close this gap. This feedback information can be seen as having different depth of elaborated information or 'specificity', including identification of areas/strengths, identification of error/mistakes/weaknesses, and suggestions for improvement (Chen, Wei, Wu, & Uden, 2009; Mory, 2004; Narciss, 2008). For example, in a study on unmediated (without prior training or support) peer assessment, seventh graders' ability to judge the performance of a peer was measured by positive and negative feedback provided (Tsivitanidou, Zacharia, & Hovardas, 2011). The participants were also found to suggest changes or revisions, which the authors commented as contributing to the 'structural components of a constructive feedback' (p.516). However, the validity and reliability of the students' assessment skills in this area were low compared to a group of expert reviewers and explicit training was recommended to improve these skills. In a comparative study of peer and teacher feedback of seventh graders within a web-based learning platform, Hovardas, Tsivitanidou & Zacharia (2014) conceptualised peer feedback quality as composed of quantitative feedback - the validity and reliability of assigned grades and qualitative feedback - positive/negative judgement and suggested changes for revisions, consistency between judgement and suggested changes, and the accuracy of content. Peer

assessors were found to be similar with expert assessors in terms of judgement and suggestions for changes, but showed lower validity and reliability in grading as well as placing less emphasis on skills needed for task completion and fewer reasoning and suggestions for change to accompany negative judgements. Further research in ways to support the skills of peer assessors were proposed, such as introducing computer-based scaffolding and training.

Survey studies of conceptions of feedback by New Zealand primary and secondary students have measured the factors on Comments for improvement, Interpersonal feedback and Negative feedback and demonstrated that students predominantly favour teacher-led feedback practices (Harris, Brown & Harnett, 2014). While this may reflect the ecological reality in New Zealand schools, the students are positive about the use of feedback for their learning and there is evidence of primary students accepting peer- and self-feedback as legitimate sources of feedback. The lack of acceptance of peer feedback for secondary students suggests that this area is in need of further research.

**Feedback levels.** Another perspective on peer feedback builds on the notion of reducing the ‘discrepancies between current understanding and performance’ by focusing on the learner’s engagement with the feedback information at the task, process, self-regulation, and self levels (Hattie & Timperley, 2007). In Hattie and Timperley’s model, interventions involving feedback are likely to be more effective when the learner’s attention is drawn to cognitive outcomes related to the task, task processing strategies, and the self-regulation strategies adopted, rather than focusing on the self.

First, Hattie and Timperley (2007) postulated that feedback can engage learners at the task level, such as providing information on correct response. Second, feedback can be aimed at the process level, such as providing task processing strategies and cues for information searching. The third level of feedback is focused on self-regulation, including the skills of self-evaluation, expanding effort in task engagement or seeking further feedback information. The fourth level of feedback is seen as directed to the “self”, usually involving praise. Examples of such feedback include, “You have done well!”, “Keep up the good work!” In nearly all learning situations, praise does not provide information on how to improve performance on the task (Kluger & DeNisi, 1996). Praise is rarely effective when students’ attention is drawn to the self, and may even have negative consequences, such as distracting the learner from the task, and encouraging effort avoidance behaviour in order to minimise

the risk to the self (Black & Wiliam, 1998a; Hyland & Hyland, 2006). Studies comparing general, non-targeted praise with ability and effort feedback have consistently indicated that general praise is not effective for learning, while suggesting caution that ability and effort feedback are contingent to other variables such as maturity of students to interpret feedback, the context and specificity of feedback and students' relationship with the teacher (Henderlong, & Lepper, 2002; Burnett & Mandel, 2010). The feedback message should focus explicitly on task, process and self-regulation aspects of learning, and thus, directly involving the learner in the knowledge building process.

Hattie and Timperley's (2007) four levels of feedback provide a potentially useful model for classifying peer feedback that could help teachers and students to articulate feedback features in a common language, and to visualise the process of feedback exchange as moving progressively towards higher quality work (Authors, 2011). At present, few studies have examined the usefulness of this model to understand peer feedback features and skills in the classrooms. In a recent study, primary and secondary teachers completed a survey questionnaire that required them to indicate their level of agreement to statements about the four feedback levels (Brown, Harris, & Harnett, 2012). Most teachers found task level feedback agreeable, with moderate agreement to process and self-regulation level feedback. Teachers only slightly to moderately agreed with the self level feedback. Although teachers were able to differentiate between the four feedback levels, the authors found that their conceptualisation of these levels differed considerably from the definitions found in the original model.

### **Prompting peer feedback**

A common approach to support targeted peer feedback is to provide students with feedback prompts (Chen, Wei, Wu, & Uden, 2009). Prompts are generally regarded as guiding questions, sentence openers, or question stems which provide cues, hints, suggestions, and reminders to help students complete a task. In general, most peer learning situations involve prompts which were specifically designed for use by students to engage their peers in collaborative discourse (e.g., Palincsar & Brown, 1984; Davis & Linn, 2003). The above review also showed that prompts may be beneficial as instructional support to guide and scaffold learners in formulating peer feedback that promotes procedural as well as deeper levels of engagement.

In peer feedback research, prompts are commonly used to support the provision of meaningful feedback that results in task completion or follow-up revision activities (Gielen et al., 2010). The need for students to provide relevant and useful feedback means that prompts are seen as serving a predominantly functional role and may be incorporated into a rubric (Cho & MacArthur, 2010), a step-by-step strategy tool (Van Steendam et al., 2010), or detailed guidance sheet (Min, 2006). In a study on improving the effectiveness of peer feedback of secondary students for learning, feedback forms with question prompts were used to cue students in providing comments on strengths, weaknesses, and justifications of writing assignments and to formulate reflections on how they used the received-feedback as well as learning experiences from the feedback exchange (Gielen et al., 2010). The increased comments on justification were found to improve writing performance but reflections did not lead to learning gains. Although it is acknowledged that there is a link between prompts and quality of peer feedback, this relationship is more implicit than explicit. It follows that the questions surrounding the prompting of peer feedback may warrant on-going scrutiny.

### **Aims of the present study – Hypotheses**

In this study, we investigated the use of question prompts to elicit reviewers' feedback to their peers, and whether the prompts help to support the posing of more comments on the use of criteria, feedback specificity and feedback levels. The prompts took the form of probing questions that require students to provide feedback on their peer's work about what they have done well, what they did poorly and how best to improve their work. In particular, prompted/unprompted peer feedback is analysed in relation to four research questions:

- (a) What are the feedback features present in reviewers' prompted and unprompted peer feedback in relation to use of criteria, feedback specificity, and feedback levels?
- (b) What are the effects of prompts on posing peer feedback in relation to use of criteria, feedback specificity, and feedback levels?
- (c) How do reviewers' draft report writing influence their given-feedback?
- (d) Do giving more specificity or level peer feedback influence report writing quality of reviewers?

The following hypotheses were tested:

- (a) The use of criteria, feedback specificity and feedback levels will be present in reviewers' feedback statements but will vary between prompted and unprompted conditions (Hypothesis 1).
- (b) Reviewers in the prompted condition will pose more peer feedback with the use of criteria, specificity (Knowledge of correct response, Knowledge of errors and Suggestions for improvement), and feedback levels (Task, Process, Self-regulation and Praise), compared to reviewers in the unprompted condition (Hypothesis 2).
- (c) Reviewers with better draft report writing skills will generate more given-feedback (Hypothesis 3).
- (d) Feedback specificity and feedback level will mediate the influence of prompts on report writing quality of reviewers (Hypothesis 4).

## **Method**

### **Participants**

Six classes of Year 12 chemistry students (16–17 year olds) from three New Zealand urban secondary schools (offering the National Certificate of Educational Achievement NCEA<sup>1</sup> chemistry curriculum) participated in this study (N = 121; 75 females and 46 males). The criteria for selecting schools include having Year 12 chemistry classes, availability of laboratory resources, and that the students had already been taught the topic on 'rate of reaction' in the same year of the study. All the participants had completed the topic on 'rate of reaction' and had experience with practical work involving planning and implementing experiments.

### **Research design and procedure**

The study adopted a quasi-experimental repeated measures pretest-treatment-posttest design. The six intact classes were randomly assigned to one of the two treatment conditions, with four classes in the prompted peer feedback condition (experimental group, N=77), and two classes in the unprompted peer feedback condition (control group, N=44). The unequal sample sizes was due to constraints of time-tabling in one school and access to practical equipment and thus, unrelated to the experimental conditions. A prior knowledge pretest for

all participants found no significant difference between the two experimental groups,  $t(114) = 1.325$ ,  $p = 0.188$ , and the assumption of homogeneity of variance has not been violated (Levene's Test for Equality of Variances is not significant,  $F = 2.382$ ,  $p = 0.125$ ).

Prior to the intervention, the teacher explained the use of the investigation report booklet which included two types of review forms – prompted and unprompted. Students in the prompted condition reviewed their peer's report using the three question prompts in the review form (adapted from Gielen, 2007):

- What did he/she do well? Give explanations to support your feedback.
- What didn't he/she do well? Give explanations to support your feedback.
- How can he/she improve on the current piece of work? Give explanations to support your feedback.

Students in the unprompted condition wrote comments to their peers using the review form without the question prompts.

The intervention consisted of two review cycles: the first cycle involved students conducting an individual experimental trial and writing a draft plan of the investigation, followed by reciprocal peer feedback and revision of their report. The second cycle began when students used their revised plan to implement their experiments, collect and analyse their data, and write a full laboratory report. This was followed by another round of peer feedback exchange and revision before final submission of their report. The class teacher for each class assisted with obtaining the practical materials and conducting the investigative task, while the researcher acted as observer. Each class was allocated four lessons to carry out the research study. The pairing of students was carried out by the class teacher, who allocated the partnership based on high and low ability pairing using previous test scores as well as observations of prior encounters of the students' group work in the laboratory.

## Measures

**Investigation report performance.** The investigative task in this study was designed based on the NCEA Level 1 chemistry unit standard<sup>2</sup> 'Investigate factors that affect the rate of a chemical reaction'. Chemical kinetics is a fundamental topic in almost every secondary school chemistry syllabus because of its importance in the understanding of basic concepts in

chemical reaction processes. Students were required to plan, implement, and write a report of a semi-open investigation on the rate of a reaction using the report booklet provided by the researcher. The report format consists of two sections, a) experimental trial and planning and b) implementation, results and evaluation. For each section, students wrote a draft and then revised this draft based on the peer feedback received. Marking of the reports was carried out by one of the researcher, and validated by a subject expert. The revised report writing performance ranged from 4.43 to 7.77 ( $M = 5.79$ ,  $SD = 1.34$ ) in the first review cycle and from 4.65 to 8.86 ( $M = 6.22$ ,  $SD = 2.14$ ) in the second review.

**Peer feedback coding and analysis.** The unit of analysis was each individual student's written peer feedback, which was identified as a self-contained statement focusing on a particular issue of report writing. Peer feedback statements were coded on the specificity of feedback and the levels of feedback. An analysis of the peer feedback statements on the use of criteria was also carried out with a scoring rubric.

**The use of criteria in peer feedback.** In this study, the criteria for use in peer feedback were developed based on the work by Robert, Gott and their colleagues (2003) on 'Concepts of Evidence' (CoE). Investigations that require students to design, implement and evaluate experiments have been recognised as playing an important role in 'science as practice' and developing 'inquiry' skills that are reflective of scientific literacy (Abd-El-Khalick et al., 2004; Duschl, Schweingruber, & Shouse, 2006). Learning science involves both an understanding of the substantive content knowledge as well as the procedures of science. There is a need of a different way to conceptualise the procedural component of the science curriculum. The skills perspective, which emphasised developing 'process skills', should be complemented by an understanding of ideas about evidence. This procedural component requires the learner to construct meaning, specifically about validity and reliability, from specific ideas about evidence; which is seen as an integral part of science and that can then be learned, understood and applied, rather than a set of skills that develop implicitly by practice (Gott & Roberts, 2008). A CoE list was developed to 'serve as a domain specification of ideas necessary for procedural understanding' (Gott & Roberts, 2003). Rather than seeing this list as isolated and routinized procedure, it can be a 'toolkit of ideas' that help students to apply and understand procedural knowledge when planning and carrying out practical investigations. A scoring rubric based on CoE was developed to analyse the participants' use of these criteria in their written peer feedback responses (Table

1). Five key concepts of evidence relevant to the investigative task in this study were selected from the CoE list — Variable structure, Fair-testing, Choosing values, Patterns and relationship in data, and Reliability and validity of data/design. The written feedback statements of each reviewer were scored on the five criteria; with 0 for not using the criterion, 1 for stating the criterion and 2 for elaborating on the criterion. Interrater reliability between the researcher and an independent rater was acceptable,  $r = 0.70$ ,  $p < 0.01$ .

---

Insert Table 1 about here

---

**Specificity of peer feedback.** The coding scheme for the specificity of feedback was developed based on feedback research which identified three common feedback specificity in the classroom — Knowledge of correct response, Knowledge of errors and Suggestions for improvement (Mory, 2004; Narciss, 2008).

**Levels of peer feedback.** The coding scheme for feedback levels was developed from Hattie and Timperley's (2007) feedback model – Task level feedback, Process level feedback, Self-regulation level feedback, Praise and Others (see Table 2).

---

Insert Table 2 about here

---

Trustworthiness of this research study was a key concern when working on the coding scheme, analysing the results and interpreting the findings (Lincoln and Guba 1985). The researcher, working with a colleague, carried out frequent discussions and debriefs for each of the coding and scoring processes to triangulate their interpretations, resolve inconsistencies and refine the coding and rating scheme (Braun and Clarke 2006; Thomas 2006). Besides establishing credibility, an internal audit was carried out with the help of a subject expert in the field of chemistry education on the dependability of the findings. To establish inter-coder reliability, the peer feedback responses of all participants were coded independently by two different coders using the final coding scheme. Cohen's kappa for peer feedback specificity and levels was at least 0.80 for all analyses.

## Results

What feedback features are present in reviewers' prompted and unprompted peer feedback in relation to use of criteria, feedback specificity, and feedback levels?

The use of criteria in peer feedback statements in both prompted and unprompted conditions were analysed by using the concept of evidence rubric (Table 3). Overall, students in both conditions were observed to be using CoE in their peer feedback. Reviewers' in the prompted condition showed a wide variation in the use of CoE ( $M = 2.45$ ,  $SD = 2.03$ ) compared to reviewers in the unprompted condition ( $M = 2.30$ ,  $SD = 1.39$ ).

---

Insert Table 3 about here

---

A closer look at the breakdown of the five CoE indicated that reviewers in the prompted condition generated peer feedback that ranged from a failure to use particular evidence (such as Choosing value' and Reliability & validity of data/design) to the frequent use of 'Fair testing' (Table 4). Reviewers in the unprompted condition were able to provide peer feedback that includes all the five criteria; again, with Fair testing as the frequently used CoE.

---

Insert Table 4 about here

---

When comparing the specificity of peer feedback, reviewers in the prompted condition were able to generate more peer feedback that focused on knowledge of errors ( $M = 3.19$ ,  $SD = 1.76$ ) and suggestions for improvement ( $M = 3.27$ ,  $SD = 2.09$ ) (Table 3). In contrast, reviewers using a generic feedback form without prompts managed to generate fewer feedback statements on Knowledge of errors ( $M = 1.95$ ,  $SD = 1.77$ ) and Suggestions for improvement ( $M = 1.07$ ,  $SD = 1.52$ ). Reviewers tended to give more knowledge of correct responses in unprompted peer feedback ( $M = 5.59$ ,  $SD = 3.54$ ) than in prompted peer feedback ( $M = 5.14$ ,  $SD = 2.91$ ), which was the predominant type of feedback response that students were familiar with.

Reviewers in the prompted condition generated more task ( $M = 9.27$ ,  $SD = 5.12$ ), process ( $M = 1.29$ ,  $SD = 1.61$ ) and self-regulation level ( $M = 0.06$ ,  $SD = 0.25$ ) feedback than reviewers in the unprompted feedback condition (task:  $M = 7.32$ ,  $SD = 3.95$ ; process;  $M = 0.34$ ,  $SD = 0.61$ ; self-regulatory:  $M = 0.02$ ,  $SD = 0.15$ ) (Table 3). Reviewers were more likely to use praise in unprompted peer feedback ( $M = 0.84$ ,  $SD = 1.38$ ) than in prompted peer feedback ( $M = 0.45$ ,  $SD = 0.77$ ). The disaggregation of feedback levels further suggested that task level feedback was the focus of most reviewers' comments, followed by

the use of process level feedback (Table 5). Only a small number of reviewers' comments were generated at the self-regulatory level.

---

Insert Table 5 about here

---

What are the effects of prompts on posing peer feedback in relation to use of criteria, feedback specificity, and feedback levels?

To compare the use of criteria, peer feedback specificity and feedback levels with the experimental conditions, a MANOVA was conducted. The multivariate main effect of prompting was significant, Wilks' Lambda = 0.619,  $F(8, 112) = 8.611$ ,  $p < .001$ , partial  $\eta^2 = .381$ . Follow up ANOVA results revealed a significant effect for the use of prompts on peer feedback specificity: Knowledge of errors,  $F(1, 119) = 13.835$ ,  $p < .001$ , partial  $\eta^2 = .104$  and Suggestions for improvement,  $F(1, 119) = 37.479$ ,  $p < .001$ , partial  $\eta^2 = .240$ , but showed no effects for Knowledge of correct response. For feedback levels, only reviewers' Process level feedback was significantly different between the two conditions,  $F(1, 119) = 14.056$ ,  $p < .001$ , partial  $\eta^2 = .106$ . No significant difference was found for the use of criteria in reviewers' comments. All values were compared at a Bonferroni adjusted alpha level =  $0.05/9 = 0.005$ . The results indicated that reviewers in the prompted condition made more comments on identifying errors, suggesting ways to help improve their peer's report, and focusing their attention on process level feedback.

How do reviewers' draft report writing influence their given-feedback?

A MANCOVA was carried out with draft report writing as covariate to explore its influence on the experimental outcomes. The draft report writing skills of reviewers were based on their performance scores on the written drafts for the two review cycles, ( $M = 10.58$ ,  $SD = 2.99$ ). Prior to MANCOVA, correlation analyses between draft report scores and feedback features were found to be reasonable except for Praise and Others ( $r = .20$  to  $.45$ ) (Table 6); these two feedback features were retained in subsequent analysis as they form part of the feedback level model. T-test between covariate and the two experimental conditions showed that there was a significant difference in draft report scores,  $t(119) = 3.982$ ,  $p < .001$  which suggested that draft report writing skills may confound the main outcome of prompting. When draft report scores were added as covariate in a MANCOVA, the multivariate outcome is less significant, Wilks' Lambda = 6.922,  $F(8, 111) = 8.611$ ,  $p < .001$ , partial  $\eta^2 = .333$ ,

further indicating that reviewers' reporting writing skills play a part in generating different feedback specificity and levels. Post hoc (Bonferroni) analyses of the univariate outcomes (adjusted for draft report scores and alpha level =  $0.05/9 = 0.005$ ) showed that the effect of prompting was only significant for Suggestions for improvement,  $F(1, 118) = 22.478$ ,  $p < .001$ , partial  $\eta^2 = .160$ .

---

Insert Table 6 about here

---

Do giving more specificity or level peer feedback influence report writing quality of reviewers?

A mediation analysis was carried out to test the hypothesis that peer feedback specificity or feedback levels may act as mediators to the relationship between prompting peer feedback and report writing quality of reviewers. Report quality was taken from a measure of the difference between final report score and draft report score, summed for both review cycles. We noted that this measure only served as a proxy for the writing quality of reviewers, as the peer feedback process is reciprocal, i.e. the reviewers may benefit from giving comments as well as the received comments. Knowledge of errors, Suggestions for improvement and Process level feedback were tested as mediators due to their significant effects in the experimental conditions. A series of linear regressions and Sobel  $z$ -tests (1982) were carried out for each mediator (MacKinnon, 2007). A reduction in regression coefficient when controlling for each of the mediators suggested that partial mediating effects were present for all three mediators. However, significant mediating effect was confirmed for Knowledge of errors,  $z = -2.94$ ,  $p = 0.003$ , and Suggestions for improvement,  $z = -3.43$ ,  $p < 0.001$  but not for Process level feedback,  $z = -1.71$ ,  $p = 0.086$ .

### Discussion

The present study examined the effects of prompting on reviewers' written comments. In particular, peer feedback of secondary students in prompted and unprompted conditions were examined to identify and compare feedback features related to the use of criteria, feedback specificity and feedback levels. Reviewers' initial writing skills and overall report writing performance were also examined to understand how these variables may play a part in giving peer feedback.

Reviewers in both experimental conditions included concepts of evidence in their comments, with prompted condition showing greater variation in the use of these criteria compared to the unprompted condition. The effect of prompting did not result in more comments with the use of criteria and a closer look at individual classes suggested that not all students are proficient in posing feedback with criteria. On one hand, the concepts of evidence rubric were found to be useful in identifying the ideas related to the five key evidence used by students during the peer feedback process; on the other, it showed that students tend to concentrate solely on ensuring a fair testing in designing experiments. This finding concurred with previous studies on the opportunities of ‘doing science’ in New Zealand classrooms (Barker, 1999; Haigh, France, & Forret, 2005; Hodson, 2009), which found that students were more proficient at carrying out specified practical tasks than the independent problem-solving tasks that were involved in open planned investigative activities. Thus, the use of prompted peer feedback exchanges in practical tasks may open up avenues for students to develop evidence ideas related to other aspects such as identifying patterns and relationships as well as concepts of reliability and validity when working on data or experimental design (e.g. Wilson, Perry, Anderson & Grosshandler, 2012).

In terms of the reviewers’ feedback specificity, it was found that the prompted condition resulted in more Knowledge of errors and Suggestions for improvement, but not on Knowledge of correct response. This finding is significant on two accounts. First, it demonstrated that secondary students are capable of posing written peer feedback with greater specificity that went beyond just giving the right answers. The provision of explicit solutions and ways of improving their work has been shown to be an effective feature of peer feedback. For example, Nelson and Schunn (2009) found that peer feedback was more likely to be implemented when a solution was provided, attributing the effect to helping the learner with missing information that are needed for task completion. Second, consistent with other studies (Cho & MacArthur, 2010; Gielen et al., 2010; Min, 2006), the successful use of question prompts suggested that this approach provided a useful tool for teachers to help young reviewers in posing more elaborate peer feedback during peer review sessions. One possible reason for the effectiveness of prompting peer feedback is that prompts help to overcome superficial processing of learning strategies by the learner by acting as “strategy activators” (Berthold, Nuckles, & Renkl, 2007; Davis & Linn, 2000; Lin & Lehman, 1999). The question prompts cued the students to draw comparisons between the task criteria and the work done (peer’s written report) by evaluating what was or was not done well. By

articulating the criteria in relation to the work done in the form of feedback to their peers, students established a standard for their work (Orsmond, Merry, & Reiling, 2002). At the same time, students have to come up with suggestions for improvement which required them to think of possible ways of ‘closing’ the gap by revising their partner’s work. In short, the prompts made the steps on identify learning gaps in relation to criteria and how to carry out revisions explicit to the learners.

This study adopted Hattie and Timperley’s (2007) model to characterise peer feedback as consisting of Task, Process, Self-regulation, Praise and Others. Interestingly, the four feedback levels were present in both conditions, with the prompted condition generating more Task, Process and Self-regulation feedback. By contrast, reviewers in the unprompted condition used more Praise in their comments. Consistent with the teacher survey study by Brown et al. (2012), Task level feedback was the most commonly used feedback whereas self-regulation feedback was rarely found in the reviewers’ comments. The effects of prompting was only significant for Process level feedback, and this further suggested that while the model provided a useful classification of peer feedback, the posing of feedback at the different levels required explicit instructional support, especially for young reviewers. This support may include explicit instructions on how to give and receive peer feedback at each feedback level (Author, 2011).

The coding scheme for feedback levels was developed in this study by operationalizing Hattie and Timperley’s feedback model. While this has helped in the coding of peer feedback statements in terms of task, process, and self levels, it raised doubts about how to conceptualise self-regulation as a coding for peer feedback statements. The notion of self-regulation encompasses a wider conception of a dynamic learning process involving executive control over one’s own learning and influenced by a multitude of factors, such as personal characteristics, social circumstances and learning conditions (see Boekaerts, 2006). The peer feedback statements by young learners would probably not fall under this category and thus, more research into this area is need to identify and conceptualise self-regulation feedback to refine the model.

When we considered the influence of draft report writing skills of the reviewers, the effect of prompting alone may not be responsible for the specificity and levels of feedback. The confounding effect of students’ draft report writing skills suggested that not all students have the necessary report writing skills to benefit from the feedback that they received. This

finding is unsurprising, given that these young reviewers had limited exposure to investigation report writing (see comment above on using CoE) and at the same time, had to grapple with posing feedback, interpreting the received feedback and purposefully using the comments to improve their reports. For the young reviewers, a combination of guided discussion on the peer feedback and prompting may be more effective (Min, 2006).

Testing the mediating effects of comments on Knowledge of errors, Suggestions for improvement and Process level feedback provided further insights on students report writing. We observed the significant mediating effect of more Knowledge of errors and Suggestions for improvement feedback, revealing that supporting peer feedback specificity improves report writing of reviewers (Cho & Cho, 2011). Although Process level feedback did not mediate reviewers report writing as a separate variable, comments at the process level draw students' attention to procedural strategies which may complement feedback specificity to improve students' reports. This interactive nature of feedback features opens up avenues for future research.

This research contributes to collaborative learning in the chemistry laboratory by addressing the need to create opportunities for students to engage in meaningful peer feedback discourse. While supporting and developing peer feedback skills and strategies are important, teachers and educators should also realise that students need to be provided with a feedback "space", where errors or mistakes are seen as learning points, and peer feedback goes beyond a corrective function to one which promotes critical discourse (Author, 2014). This implies that we need to see peer feedback as negotiating meaning and connecting ideas, rather than providing the "right" answers. Instead of having peer feedback sessions that focus solely on marking and providing confirmatory responses, students should be encouraged to discuss the feedback, ask further questions for clarification and share alternative viewpoints. The feedback model, with elements of feedback levels, can be seen as an instructional strategy to promote active critical engagement with scientific text and issues. Here, the advantage is seen as providing opportunities to explore and use the language of science in interpreting, evaluating, justifying, and reviewing the written science reports of their peers and then finding ways of clearly communicating that feedback. Giving and/or receiving peer feedback mirrors the important practice of 'peer reviewing' in the scientific community and thus, facilitates the inclusion of ideas related to the nature of science, as well concepts of evidence.

Acknowledging that the results from this intervention study were limited in generalisability due to the non-random selection of the treatment groups and the short duration of the intervention, the findings were positive in suggesting that question prompts guided students in formulating peer feedback that directed their peers to the learning gaps and indicated how best to improve their performance. The findings should be interpreted in light of the unequal sample size, although precautions were taken to ensure equivalence of both experimental groups prior to the start of the intervention, such as prior knowledge testing and Levene's test of sample variances. The mediation analysis on peer feedback features as mediators suggested that more research can be done to investigate the relationship between instructional support, peer feedback quality and report writing outcomes as well as the possible nested nature of reviewers' peer feedback in different classes within the participating schools.

## References

- Abd-El-Khalick, F., Boujaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A., et al. (2004). Inquiry in science education: International perspective. *Science Education*, 88(3), 397–419.
- Althausen, R., & Darnall, K. (2001). Enhancing critical reading and writing through peer reviews: An exploration of assisted performance. *Teaching Sociology*, 29, 23–35.
- Andrade, H. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. Andrade & G. Cizek (Eds.), *Handbook of formative assessment*. New York: Routledge.
- Author (2011). [Details removed for peer review]
- Authors (2011). [Details removed for peer review]
- Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–73.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Berkshire, England: Open University Press.
- Berthold, K., Nuckles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction*, 17, 564–577.
- Boekaerts, M. (2006). Self-regulation and effort investment. In K.A. Renninger & I.E. Sigel (Eds.), *Handbook of child psychology: Vol.4. Child psychology in practice* (pp. 345-377). New York: Wiley.
- Boud, D. (1995). *Enhancing learning through self-assessment*. London: Kogan Page.
- Boud, D., Cohen, R., & Sampson, J. (1999). Peer learning and assessment. *Assessment & Evaluation in Higher Education*, 24(4), 413.
- Brown, G.T.L., Harris, L.R., & Harnett, J. (2012). Teacher beliefs about feedback within an assessment for learning environment: Endorsement of improved learning over student well-being. *Teaching and Teacher Education*, 28, 968-978.

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Burnett, P.C., & Mandel, V. (2010). Praise and Feedback in the Primary Classroom: Teachers' and Students' Perspectives. *Australian Journal of Educational & Developmental Psychology*, 10, 145-154.
- Chen, N.-S., Wei, C.-W., Wu, K.-T., & Uden, L. (2009). Effects of high level prompts and peer assessment on online learners' reflection levels. *Computers & Education*, 52(2), 283-291.
- Cho, Y.H., & Cho, K. (2011). Peer reviewers learn from giving comments. *Instructional Science*, 39, 629–643.
- Cho, K., & MacArthur, C. (2010). Student revision with peer and expert reviewing. *Learning and Instruction*, 20(4), 328–338.
- Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education*, 48(3), 409–426.
- Davis, E. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *The Journal of the Learning Sciences*, 12, 91–142.
- Davis, E. A., & Linn, M. C. (2000). Scaffolding students' knowledge integration: Prompting for reflection in KIE. *International Journal of Science Education*, 22, 819–837.
- Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2006). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: The National Academies Press.
- Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70, 287–322.
- Falchikov, N. (2001). *Learning together: Peer tutoring in higher education*. London & New York: RoutledgeFalmer.

- Falchikov, N. (2005). *Improving assessment through student involvement: practical solutions for learning in higher and further education*. New York: RoutledgeFalmer.
- Frederiksen, J. R. & White, B. J. (1997). Reflective assessment of students' research within an inquiry-based middle school science curriculum. In Proceedings of Annual meeting of the AERA conference, vol . Chicago, IL.
- Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning & Teaching in Higher Education*, 1, 3–31.
- Gielen, S. (2007). *Peer assessment as a tool for learning*. Unpublished doctoral dissertation, Leuven University, Leuven, Belgium.
- Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction*, 20(4), 304–315.
- Gott, R. & Duggan, S. (2003). *Understanding and using scientific evidence*. London: Sage.
- Gott, R., & Roberts, R. (2008). *Concepts of evidence and their role in open-ended practical investigations and scientific literacy*; background to published papers. UK: The School of Education, Durham University.
- Haigh, M., France, B., & Forret, M. (2005). Is “doing science” in New Zealand classrooms an expression of scientific inquiry? *International Journal of Science Education*, 27, 215–226.
- Harris, L.R., Brown, G.T.L., & Harnett, J. (in Press). Understanding classroom feedback practices: A study of New Zealand student experiences, perceptions, and emotional responses. *Educational Assessment, Evaluation and Accountability*.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Henderlong, J. & Lepper, M.R. (2002). The Effects of Praise on Children's Intrinsic Motivation: A Review and Synthesis. *Psychological Bulletin*, 128(5), 774–795.
- Hodson, D. (2009). *Teaching and learning about science: Language, theories, methods, history, traditions and values*. Rotterdam, Netherlands: Sense.

- Hovardas, T., Tsivitanidou, O.E., & Zacharia, Z.C. (2014). Peer versus expert feedback: An investigation of the quality of peer feedback among secondary school students. *Computers & Education, 71*, 133–152.
- Hyland, K., & Hyland, F. (Eds.). (2006). *Feedback in second language writing: Contexts and issues*. Cambridge, UK: Cambridge University Press.
- Kim, M. (2009). The impact of an elaborated assessee's role in peer assessment. *Assessment & Evaluation in Higher Education, 34*(1), 105-114.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284.
- Lin, X., & Lehman, J. D. (1999). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching, 36*, 837–858.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.
- Lockhart, C., & Ng, P. (1995). Analyzing talk in ESL peer response groups: stances, functions and content. *Language Learning, 45*, 605–655.
- Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing, 18*(1), 30–43.
- MacKinnon, D.P. (2007). *Introduction to Statistical Mediation Analysis*. Mahwah, NJ: Erlbaum.
- McConlogue, T. (2014). Making judgements: investigating the process of composing and receiving peer feedback. *Studies in Higher Education, 1*-13.
- Min, H. T. (2005). Training students to become successful peer reviewers. *System, 33*, 293–308.
- Min, H. T. (2006). The effects of training peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing, 15*(2), 118–141.

- Mory, E. H. (2004). Feedback research review. In D. Jonassen (Ed), *Handbook of research on educational communications and technology* (pp.745–783). Mahwah, NJ: Lawrence Erlbaum Associates.
- Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. J. G. Van Merriënboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3<sup>rd</sup> ed., pp.125–143). Mahwah, NJ: Lawrence Erlbaum Associates.
- Nelson, M., & Schunn, C. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science*, 37(4), 375–401.
- Ngar-Fun, L., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education*, 11(3), 279–290.
- Orsmond, P., Merry, S., & Reiling, K. (2000). The use of student derived marking criteria in peer and self assessment. *Assessment and Evaluation in Higher Education*, 25(1), 23–38.
- Orsmond, P., Merry, S., & Reiling, K. (2002). The use of exemplars and formative feedback when using student-derived marking criteria in peer and self assessment. *Assessment and Evaluation in Higher Education*, 27(4), 309–323.
- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension fostering and comprehension monitoring activities. *Cognition and Instruction*, 1(2), 117–175.
- Ploegh, K., Tillema, H. H., & Segers, M. S. R. (2009). In search of quality criteria in peer assessment practices. *Studies in Educational Evaluation*, 35, 102-109.
- Prins, F., Sluijsmans, D., & Kirschner, P. (2006). Feedback for general practitioners in training: Quality, styles and preferences. *Advances in Health Sciences Education*, 11, 289–303.
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioural Science*, 28, 4–13.
- Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT Journal: English Language Teachers Journal*, 59(1), 23–30.

- Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119–144.
- Sadler, R. (2009). Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading in higher education. *Assessment and Evaluation in Higher Education, 34*, 159-179.
- Sobel, M.E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology, 13*, 290-312.
- Sluijsmans, D. M. A., & van Merriënboer, J. J. G. (2000). *A peer assessment model*. Heerlen: Open University of the Netherlands.
- Sluijsmans, D. M. A., Brand-Gruwel, S., van Merriënboer, J. J. G., & Bastiaens, T. J. (2002). The training of peer assessment skills to promote the development of reflection skills in teacher education. *Studies in Educational Evaluation, 29*(1), 23–42.
- Strijbos, J. W., Narciss, S., & Dünnebier, K. (2010). Peer feedback content and sender's competence level in academic writing revision tasks: Are they critical for feedback perceptions and efficiency? *Learning and Instruction, 20*(4), 291–303.
- Thomas, D. R. (2006). A general inductive approach for analysing qualitative evaluation data. *American Journal of Evaluation, 27*, 237–246.
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*, 249–276.
- Topping, K. (2005). Trends in peer learning. *Educational Psychology, 25*(6), 631–645.
- Tsivitanidou, O.E., Zacharia, Z.C., & Hovardas, T. (2011). Investigating secondary school students' unmediated peer assessment skills. *Learning and Instruction, 21*, 506-519.
- Van Steendam, E., Rijlaarsdam, G., Sercu, L., & Van den Bergh, H. (2010). The effect of instruction type and dyadic or individual emulation on the quality of higher-order peer feedback in EFL. *Learning and Instruction, 20*(4), 316–327.

Wilson, T., Perry, M., Anderson, C.J., & Grosshandler, D. (2011). Engaging young students in scientific investigations: prompting for meaningful reflection. *Instructional Science*, *40*, 19–46.

**Footnote**

<sup>1</sup> NCEA is the national qualification system for students at New Zealand secondary schools.

<sup>2</sup> <http://www.nzqa.govt.nz/qualifications-standards/qualifications/ncea/>

**Tables**

**Table 1. Scoring rubric for Concepts of Evidence in the use of criteria in peer feedback**

Concepts of Evidence in feedback	Description	0	1	2
Variable structure	Identifying and understanding the basic structure of an investigation in terms of variables and their types.	No or erroneous use.	State criteria on variables.	Elaborate on independent and dependent variables.
Fair testing	'Fair tests' aim to isolate the effect of the independent variable on the dependent variable. By changing the independent variable and keeping all the control variables constant, validity is ensured.	No or erroneous use.	State criteria on fair testing.	Elaborate on manipulating variables.
Choosing values	Making informed choices about sample, relative scale, range, interval and number of readings.	No or erroneous use.	State the criteria on values.	Elaborate on choice of values.
Patterns and relationship in data	Patterns represent the behaviour of variables so that they cannot be treated in isolation from the physical system that they represent. In this investigation, the relationship between temperature and initial rate of this reaction is linear/proportional and the pattern of association is seen as causal.	No or erroneous use.	State the relationship between variables.	Elaborate on types of relationship between variables and patterns in data.
Reliability and validity of data/design	Evaluating the whole investigation by considering the design of the investigation, ideas associated with measurement, with the presentation of the data and with the interpretation of patterns and relationships, in relation to reliability and validity of data.	No or erroneous use.	Cursory mention on reliability and validity.	Elaborate on aspects of reliability and validity.

**Table 2. Coding scheme on peer feedback specificity and levels**

Main category	Definition	Example
Specificity of peer feedback		
Knowledge of correct response	Provides correct answers or results.	He wrote the prediction accurately and provided a reason for the prediction.
Knowledge of errors	Indicates incomplete, incorrect or missing responses.	She did not include the equipment needed for this experiment in the step-by-step methods.
Suggestions for improvement	Provides solution, strategies and corrective approaches.	He could have described the steps accurately to ensure a fair test so that another person could follow the steps to repeat the experiment.
Levels of peer feedback		
Task-level (TL)	Provides information about the correctness of the learner's responses. Also informs the learner of the correct answer, but without suggesting how to revise the response. May provide indication of error/incorrect response or location of mistakes.	"She explained the limitations well but didn't really say why it was reliable and didn't refer to the data in drawing conclusions." "He wrote the predictions accurately but didn't give a reason for his prediction, which could have increased the quality of his answer."
Process-level (PL)	Provides strategies/cues/hints/examples for error detection, information search or steps to revise report. May suggest explanation or justification for correct/incorrect response and reason for the use of a particular search strategy or revision approach.	"He has clearly showed the controlled variables. He should emphasise on fair testing in the method by highlighting the use of controlled variables." "She carried out the experiment well but could've given a better evaluation on results by explaining why these results have occurred, e.g., explain why the results are not accurate."
Self-regulation level (SL)	Provides reflective or probing questions that guide the learner in self-evaluation, seeking additional information, or monitoring of learning progress.	"What would happen if you changed the gap between temperatures by a larger amount?" "Why do you think this outcome (prediction) will occur?"
Praise (P)	Remarks that are directed to the 'self' mainly to give encouragement or affirmation and contains little or no task-related information.	"You are doing great!" "Well done!" "Carried out the experiment well."
Others (OT)	Comments that are ambiguous or unrelated to the task.	"He didn't finish because he was absent." "Improve on spelling."

**Table 3. Mean and standard deviation for feedback features and report writing in prompted and unprompted conditions**

	Prompted condition (N = 77)		Unprompted condition (N = 44)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Use of criteria	2.45	2.03	2.30	1.39
Knowledge of correct response	5.14	2.91	5.59	3.54
Knowledge of errors	3.19	1.76	1.95	1.77
Suggestions for improvement	3.27	2.09	1.07	1.52
Task level feedback	9.27	5.12	7.32	3.95
Process level feedback	1.29	1.61	0.34	0.61
Self-regulation level feedback	0.06	0.25	0.02	0.15
Praise	0.45	0.77	0.84	1.38
Others	0.53	1.24	0.09	0.29
Final report writing score	13.34	3.86	9.84	2.58

**Table 4. Mean and standard deviation for each class on the use of concepts of evidence**

Classes	N	Variable structure		Fair Testing		Choosing values		Patterns & relationships in data		Reliability & validity of data/design	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Prompted											
1	14	.50	.519	.71	.825	.14	.363	.29	.469	.43	.646
3	26	.54	.706	.85	.732	.31	.549	.69	.471	.69	.471
4	15	.07	.258	.13	.352	.00	.000	.07	.258	.00	.000
6	22	.64	.658	1.09	.610	.32	.646	.59	.590	.82	.733
Unprompted											
2	23	.57	.507	.57	.507	.22	.422	1.00	.426	.61	.583
5	21	.14	.359	.62	.740	.10	.301	.43	.507	.29	.463

**Table 5. Mean and standard deviation for each class on peer feedback levels (task, process, self-regulation, Praise and Others)**

Classes	Task level feedback		Process level feedback		Self-regulation level feedback		Praise		Others	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Prompted										
1	5.93	4.25	1.21	2.05	0.14	0.36	0.07	0.27	0.00	0.00
3	11.00	4.92	1.54	1.17	0.08	0.27	0.54	0.86	0.69	1.74
4	5.93	3.94	0.00	0.00	0.00	0.00	0.47	0.52	1.27	1.22
6	11.64	4.45	1.91	1.85	0.05	0.21	0.59	0.96	0.18	0.50
Unprompted										
2	7.65	3.88	0.39	0.66	0.04	0.21	0.61	0.99	0.09	0.29
5	6.95	4.09	0.29	0.56	0.00	0.00	1.10	1.70	0.10	0.30

**Table 6. Correlations among the feedback features and draft report writing score**

Variable	1	2	3	4	5	6	7	8	9
1. Knowledge of correct response	--								
2. Knowledge of error	.28**	--							
3. Suggestions for improvement	.30**	.67**	--						
4. Task level	.72**	.73**	.72**	--					
5. Process level	.33**	.46**	.55**	.36**	--				
6. Self-regulation level	.05	.14	.32**	.14	.23*	--			
7. Praise	.38**	-.07	.02	-.01	-.01	-.02	--		
8. Others	-.02	.06	.06	-.16	-.11	-.08	.11	--	
9. Draft report writing score	.25**	.36**	.45**	.40**	.45**	.20*	-.04	-.12	--

\*\* $p < 0.01$  level.

\*  $p < 0.05$  level.