A Study of Reciprocal Peer Critiquing

Alison Gemmell

March 2008

Psychology Dept., University of Glasgow

Supervisor: Dr Steve Draper

Abstract

Reciprocal peer critiquing has been advocated for many years as being advantageous to students for a range of reasons such as improving confidence in their work, and reducing errors. Most studies, though, focus on students' appraisal of the critiquing task itself, and not on whether participating in the task improves academic performance. Students' evaluation of a reciprocal peer critiquing exercise introduced by the Department of Psychology at the University of Glasgow indicated a variety of reasons why the students found the task helpful. This study was designed to test the various individual components of the critiquing task in order to determine which aspect, if any, was most beneficial in terms of improvement in students' answers to a short critical thinking exercise. Subjects wrote three passage answers, between which the interventions were carried out. Subjects' answers were marked and despite an overall improvement across the passages for all groups, the only significant difference was found between the controls and those carrying out the full task. This suggested that in this study, no individual aspect of the peer critiquing task showed significant improvement to subjects' answers in a short critical thinking exercise, but the task as a whole did cause significant improvement.

Introduction

Students often give very positive feedback to tutors about the use of reciprocal peer critiquing exercises. Although studies, including one at the University of Glasgow, have reported many benefits of peer critiquing, few detail whether or not such tasks improve students' academic performance in the coursework the task is being used upon. Thus the purpose of this study was to carry out a preliminary investigation to determine which aspect of a peer critiquing exercise was most advantageous in improving subjects' answers, if at all.

What is Reciprocal Peer Critiquing?

In reciprocal peer critiquing, students offer each other structured critical feedback on a piece of coursework (Draper, 2007). The task is often carried out in small tutorial groups of around six to eight students, and each student critiques two or three pieces of work. In most cases, the work is critiqued using a set feedback document, advising the students on what aspects of the work they should pay particular attention to, for example the quality of the write-up. However, these key aspects will obviously differ depending on the coursework and the degree course the exercise is being employed in. In reciprocal peer critiquing, the students do not allocate a mark to each others work - this is peer assessment. Research has found that students are not keen to grade their peers' work, even more so when they are friends out-with the tutorial group, and they tend to be more lenient than teacher- and even self assessments. Despite students' lack of enthusiasm at marking their peers' work, it has been found to be "very useful as a step towards understanding and exercising standards" (Draper, 2007), and some tutors may wish to have students allocate a mark at the end of the feedback.

Support for Peer Critiquing

Many university courses use a variety of methods of assessment to mark students' work and help them to attain the course learning objectives. There is an emphasis for courses to produce students with a broad range of what are commonly referred to as transferable skills (Assiter, 1995), which are thought to promote lifelong learning.

Peer review is common practice among academics in many disciplines. It helps ensure quality in published work, and provides documented critical feedback on the work of others. The paper's content is looked at, as is its legibility and use of language. Learning to review work is taught through experience, so in order to introduce students to the process of peer review many institutions employ a peer review-type exercise such as peer critiquing which can be applied to various pieces of coursework, and can be altered to make the task suitable for almost any discipline.

Teachers have reported reciprocal peer critiquing tasks to have many benefits. Bruffee (1978) was one of the first to advocate peer critiquing, and came up with the 'Brooklyn Plan' programme for students attending a Writing Centre at Brooklyn College. His programme was carried out over several weeks, with each student writing two papers and swapping them with a peer for feedback. Each paper was critiqued four times, each time looking at a different aspect, for example, its main point (thesis), structural issues or content. Bruffee claimed that students participating in the Brooklyn Plan "learn and practice judgement collaboratively, through a progressive set of analytical and evaluative tasks applied to each other's academic writing, in a context which fosters self-esteem" (Bruffee, 1978: 450).

Modifications of the Brooklyn Plan have been designed for use by any course subject where written assignments are required. Dunn (1996) employed a peer critiquing exercise as part of a statistics course and found encouraging responses from the students, with fewer errors and increased discussion in final drafts than in previous years. Falchikov (1996) adapted Bruffee's Brooklyn Plan (1987) to fit an eight-week period. In her version, criteria for an essay were agreed upon by students and teachers. Peer critiquing was carried out in class, on the first draft of the essay. Critiquing involved summarising the paper before areas of strength were pointed out, as were areas for possible improvement. In this study, Falchikov had students write a reflective summary of their involvement both in receiving and providing feedback. Students were told that the reflective work would constitute 20% of their mark, with the other 80% allocated from the essay. After the exercise, the feedback statements were studied and it was found that students gave more positive feedback than negative. Falchikov noted that positive feedback may be encouraging to students, but that negative feedback is required for "stimulating reflection and change" (Falchikov, 2001: 275). An important finding in this study was that students are able to offer practical and constructive feedback which can often be provided sooner than a teacher would be able to, although the question remains whether or not students' essays improved due to participation in the reciprocal critiquing task.

Peer Critiquing Within the University of Glasgow Psychology Department

As part of their Psychology Honours (level 3 and 4) coursework at the University of Glasgow, students undertake Critical Reviews (CRs). Students write two CRs in their third year, and one in their fourth year. Of the two written in third year, the first is a purely formative exercise, whereas the second (along with the CR written in their fourth year) is summative. Formative assessment is beneficial to students in teaching them to understand the feedback and use it to aid their writing of summative work. Topping et al. (2000) claim that formative assessment can also support deeper learning approaches. The formative CR is marked by the student's tutor, but as he/she is marking several CRs as well as coping with their usual workload, little feedback is usually provided.

Several tutors within the Department of Psychology at the University of Glasgow provide their CR groups (a maximum of six students) with a structured feedback sheet which the students can use to critique a draft copy of their peers' work. Students are asked to point out good aspects and areas of improvement for the quality of the research, the quality of the write-up, and the quality of the critical analysis. There is also space for any other comments the student assessor may wish to make. Most tutors have each student critique the work of two peers (thereby having their own work critiqued by two different peers) in order to make sure the benefits of the task are gained. Research into students' evaluation of the peer feedback process (Morrow, 2006) has highlighted various reasons why students find the critiquing task helpful. Of the five questions asked about the helpfulness of various aspects of the exercise, the median group response was always value 4 or 5 on the likert scale, where 5 indicated 'very helpful'.

In the exercise employed by the department, there are various aspects that could potentially be of greatest benefit to the students. Firstly, the task involves reading a copy of a peer's CR, and filling in the feedback sheet. Morrow (2006) noted that 93% of students found being able to read and critique a peer's CR to be 'helpful' or 'very helpful'. Nobody who participated in her study found this to be a negative experience. Secondly, 71% of participants in Morrow's study found it helpful to get positive feedback from their peers on particularly strong areas of their work, while 79% found it helpful to have their peers point out areas of potential improvement. Morrow found that 80% of participants

felt the peer review process was helpful in improving confidence in "evaluation and assessment of CRs", and 63% felt the exercise had raised their confidence in "knowing what makes a good CR", and when asked if the peer review process had been helpful in improving their confidence in their capability to think critically, 82% of participants agreed. These results suggest that students participating in the peer critiquing process feel it is beneficial to those giving the feedback as well as those receiving it. However, this is only the *feelings* of the students participating in the exercise. It remains to be seen if peer critiquing exercises improve CR marks, and if so, is there one particular aspect of the exercise that is most beneficial, or is it the task as a whole that is so helpful?

The Present Study

This study was designed to test the individual components of a peer critiquing exercise, similar to the one presently used by some of the Department of Psychology at the University of Glasgow, in order to establish if one aspect of the task was more beneficial in improving students' answers than others, or whether it is completing the exercise as a whole that leads to improved marks. This study employed the use of a short critical thinking task as opposed to a Critical Review. This was done for several reasons, the most obvious one being the time available for the task – subjects spent 5 minutes writing each passage and a little longer on the appropriate intervention technique. This allowed several answers to be written within the space of an hour, compared with the weeks of preparation and writing time allocated for a CR. Carrying out the study using a short task instead of the full CR reduced the validity of the findings, and was harder to apply a marking scheme to, but it did allow for several attempts at answers in order to make clear any differences between groups, across the task.

Three short essay titles were chosen in the form of statements, and subjects were asked to discuss each statement with their answer in the form of a passage of around 14 lines of text. Five groups were created, each one looking at a different aspect of the full critiquing exercise. Group one was the control group, and subjects simply wrote answers for the three statements. Group two subjects wrote their passage then read sample answers for the same title. This was in order to test if simply reading examples would improve performance. Group three subjects wrote their answer which was then critiqued by a peer and they were able to read the feedback before moving on to the next title. This was designed to test if receiving feedback improves performance. Group four subjects wrote their own answer then critiqued two sample answers. This tested whether or not the process of giving feedback to other people improves own performance. Group five was the combination group. Subjects in this group carried out the full exercise of writing their own answer, reading and critiquing sample answers, as well as receiving feedback on their own answer. This was to test if the full exercise improves performance.

In each case, interventions were carried out after subjects wrote their own answer to the first title statement, before they moved on to the second title, and again before they moved on to the third title. This allowed the intervention to be tested, to see if it caused an improvement across the three attempts at the critical thinking task.

It was hypothesised that the mean scores for the intervention groups (groups two – five) would improve significantly across the three passages. It was also predicted that group five, who carried out the full peer critiquing task, would show the most improvement across the three passages compared with all other groups.

Method

Design

A mixed design was employed to look for effects of the interaction between the passage number and the group number, as well as the effect of passage number, and group number themselves. In this study, the Independent Variables are the five experimental groups, and the three passage titles. The Dependent Variables are the passage scores for each participant, and the mean group scores for each of the three passages.

Subjects

25 volunteer female third year Psychology students from the University of Glasgow were randomly allocated to one of five groups. Each group was made up of five participants, with an average age of 21 years.

Materials

Booklets were made up giving full detailed instructions of the task, which varied slightly for each group. Each subject answered three topic titles which are in appendix one. Subjects wrote directly onto the booklets. Subjects in two of the groups (groups three and five) required loose copies of the Structured Feedback Sheet. A clock with a second hand was clearly displayed, allowing subjects to monitor the precise time spent on each passage, keeping to the time limits set out in the booklets.

Procedure

The subjects were randomly assigned to a group, and the experimenter clearly read out the instructions. The five groups were as follows: group one – write only (control group - subjects wrote their own passage. They did not read any examples, or give or receive any critical feedback); group two – read only (subjects wrote their passage, then read six example passages written by other students); group three – receive critique (subjects wrote their passage which was then critiqued by a peer and they were allowed to read the feedback); group four – give critical feedback (subjects wrote their passage, then read and critiqued two example passages); group five – combination group (subjects wrote their own passage, read three example passages, critiqued two further examples, and read feedback on their own passage which was critiqued by a peer).

The instructions were set out at the start of each booklet for the subjects to refer back to at any point during the exercise. The three titles were presented in the same order for each group. The subjects worked their way through the booklet, timing themselves accurately when required, and indicated to the experimenter when they had reached the end of the task.

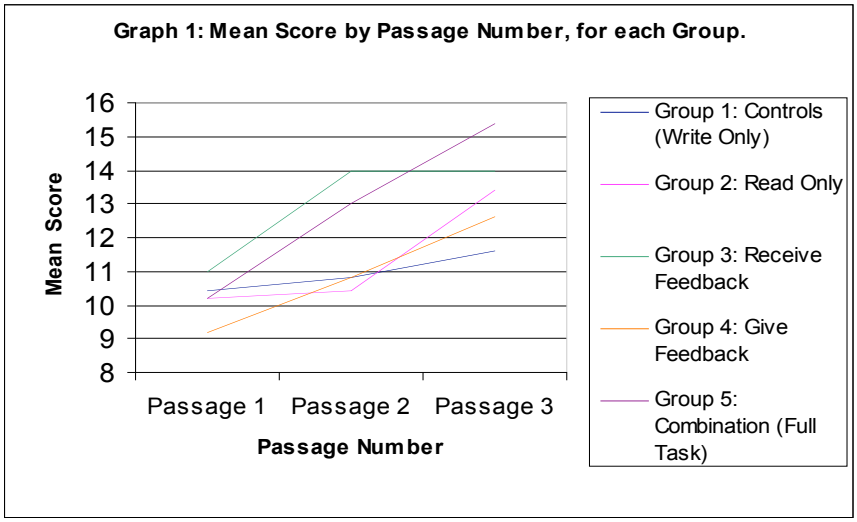
Results

Each subject's answers to the three passage titles were marked using the marking scheme in the appendix. The breakdown of each subject's score for each passage can also be found in the appendix. The mean scores were calculated, and are displayed in table 1, along with the overall improvement in mean score from passage one to passage three.

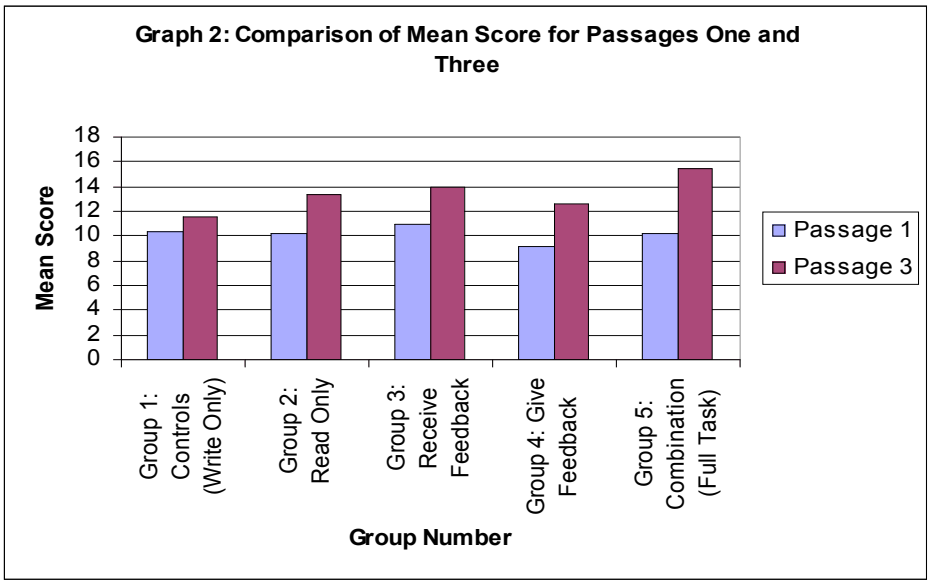
Table 1: Summary of Mean

	Passage 1	Passage 2	Passage 3	Overall Improvement
Group 1	10.4	10.8	11.6	+ 1.2
Group 2	10.2	10.4	13.4	+ 3.2
Group 3	11.0	14.0	14.0	+ 3.0
Group 4	9.2	10.8	12.6	+ 3.4
Group 5	10.2	13.0	15.4	+ 5.2

A clearer depiction of the data can be seen by looking at it in graph format.



From the table and the graph it can clearly be seen that all five groups showed an overall improvement in the mean score, between passage 1 and passage 3. Graph 2 below shows the comparison in mean scores between passage one and passage three, for the five groups. This shows the overall improvement across the task.



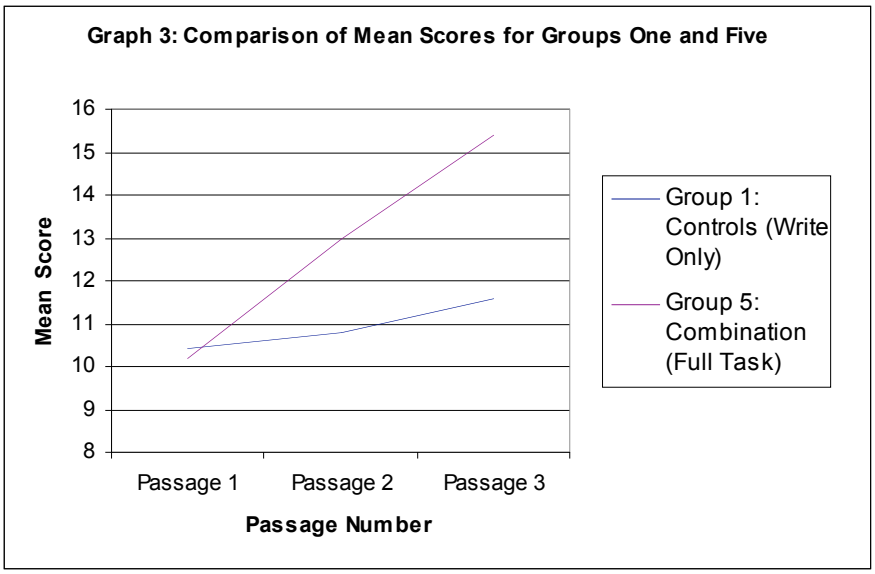
A repeated measures analysis of variance was carried out, and the results are summarised in table 2 below.

Table 2: Effects of Passage Number and Group Number

Source	SS	df	MS	F	p
Passage Number	128.000	2	64.00	29.268	0.000
Group Number	66.267	4	16.567	1.626	0.207
Passage Number*Group Number	34.533	8	4.317	1.974	0.075

The statistical results show that there was a highly significant effect of passage number ($p<0.001$). This suggests that the overall mean scores for all 25 subjects differed significantly across the three passages. The results show a non-significant result for the effect of group number ($p=0.207$), indicating that the mean scores for each group were not significantly different. A non-significant result was also found for the interaction between passage number and group number ($p=0.075$) using the repeated measures ANOVA, suggesting that each group did not differ significantly from any other across the three passages - although there is suggestion of a data trend.

It was decided to use a one way analysis of variance to compare the means for groups one and five, for passages one and three, due to the fact that this test is more powerful than the repeated measures test, and type II errors are reduced. Graph 3 shows a plot of the mean scores for all three passages for groups one and five only.



This illustrates that both the control group and the group carrying out the full critiquing task showed an improvement across the passages. The statistical results from the one way ANOVA are summarised in table 3 below.

Table 3: Comparison of Means for Groups One and Five for Passages One and Three

Source	**SS**	**df**	**MS**	**F**	**p**
Passage 1	0.100	1	0.100	0.036	0.854
Passage 3	36.100	1	36.100	14.157	0.006

From table 3 it can be seen that groups one and five did not differ significantly in the mean results for passage one ($p=0.854$) however after group five's intervention, the two groups' mean results for passage three were found to be significantly different ($p=0.006$).

Discussion

Using the repeated measures ANOVA a highly significant difference was found for passage number, indicating that the *overall* mean score improved significantly across the three passages. However, a non-significant result was found for group number, meaning that groups one to five did not differ significantly from one another, when the mean of all three passages were compared. The repeated measures ANOVA found the interaction between group number and passage number to be non-significant although the data suggested a trend, meaning that there were differences between the mean scores for each passage, for all five groups, but they were not different enough to be significant. Despite being non-significant, the results for the interaction support the fact that the intervention groups (groups two - five) showed some improvement across the three passages. It was found that group one also showed an improvement in mean score, although non-significant. This is illustrated by Graph 1. As group one were the control group, who simply wrote answers with no intervention, this would suggest that the effect of practicing the task is beneficial in slightly improving students' answers in short critical thinking tasks.

Graphs 2 and 3 both show that group five improved much more between passage one and passage three than group one, and although the repeated measures ANOVA found no significant difference between the two groups, more in-depth analysis was carried out using a one-way ANOVA comparing these two groups' mean scores for passages one and three. A one-way ANOVA is a more powerful test when comparing two groups than the repeated measures ANOVA, and it reduces the risk of type II errors (failure to reject a false null hypothesis) because the 95% confidence interval is narrower and does not span zero, as it does in the case of the repeated measures ANOVA. The one-way ANOVA found a non-significant difference ($p=0.854$) between groups one and five for passage one. This is as expected, as no intervention had been carried out at this stage, and all subjects should have performed at roughly the same level. When the means for passage three were compared, a highly significant difference was found ($p=0.006$) indicating that carrying out the full peer critiquing exercise had in fact greatly improved subjects' performance when compared with the control group.

Thus the null hypothesis could not be rejected for the first hypothesis as none of the intervention groups improved *significantly* across the three passages, though they all showed slight improvement in mean score. The null hypothesis was rejected for the second prediction as group five showed the most improvement overall, with significantly better results than the control group.

Relation to Previous Study

The findings from this study imply that all individual aspects of the reciprocal peer critiquing task cause some improvement to performance at writing answers in a short critical thinking task, although the only significant improvement was found in the group carrying out the whole task. This would suggest that in order to gain benefits to academic performance, the task as a whole should be carried out, thereby giving students the opportunity to read examples of peers' work, give feedback to peers, and receive feedback on their own work. These findings are surprising when compared with the student's evaluation of the peer critiquing task at the University of Glasgow (Morrow, 2006). Morrow found that the opportunity to read other people's CR's was rated as the most helpful aspect towards improving the students' own work. This study found that the most helpful individual aspect of the critiquing

exercise was giving feedback to peers. This intervention group improved by a mean of 3.4 from passage one to passage three compared with the group who only read examples of others' work, who showed an overall improvement of 3.2. Only 63% of respondents in Morrow's study felt that giving feedback to peers was helpful in improving their own CR, compared with the 93% who felt that reading other people's work was helpful to them.

One major criticism of Morrow's (2006) study is the number of participants involved. A sample of 15 students responded to most of the questions, with as few as 11 answering some of the questions. With such high percentages of positive evaluations from such a small sample, it is possible that only students who had felt positively about the use of the peer critiquing exercise felt motivated to respond to the questionnaire. It would be interesting to look at the attitudes and responses from a much larger sample, including the opinions of students ambivalent about the use of the task, but as the employment of the exercise is entirely voluntary and the decision of the individual tutor, this may be difficult to collate.

Limitations of this study

The sample size of 25 subjects meant only 5 subjects could be allocated to each group. It was originally hoped for more volunteer participants, but the study required level three psychology students only, and it was found they felt no motivation to take part despite advertising the potential learning benefits to them. The small sample size reduced the reliability of the findings, as one subject may have performed exceptionally better, or worse, that the others in the group which may have had a large effect on the mean score, affecting the results. The sample was also made up entirely of females. This was not intentional, and it would have been interesting to note any gender differences.

The wording of the question titles were designed to invoke a discussion of both sides of the argument, by including the instruction to "discuss". However, analysis of the subjects' answers seem to suggest that some participants either misunderstood, or could not think of arguments for both sides, as several subjects wrote answers full of opinions for one side of the topic with no mention of possible reasons for the other. As the exercise was on such a small scale, with only short passages being marked, giving arguments for only one side greatly reduced the score the subject received for that passage. This could potentially have had a huge effect on the results.

The marking scheme employed in this study was based upon the criteria used in the peer feedback form. It was decided these would be suitable measurements for testing improvement across the task. Difficulty was found when trying to apply scores to the answers, and it was felt that a marking scheme such as the one being used would be more successful on a longer piece of work. In order to increase the consistency in scoring such short pieces of work, only one marker was used to score all subjects answers.

The present study did not take into account the feedback that was given by the subjects. Morrow (2006) differentiated between the helpfulness of receiving positive comments, and comments on areas of improvement. She found that 71% of respondents found it helpful to receive comments about what was good, and 79% found it helpful to receive comments on areas of improvement. It may have been interesting to see whether those who received more comments

on areas of improvement improved to a greater extent that those who received mainly comments praising their efforts. Nor did this study look at anything other than improvement at the task. Students' opinions on what, if any, other benefits they feel they gained would have been another aspect to take into account. These could have been compared with the results of actual improvement, to investigate whether the students perceived the most beneficial aspect in this short task.

Implications of this Study

In the present study, the control group also showed an improvement, albeit the smallest improvement compared with the other four groups. This suggests that there must be a simple effect of practice involved. Although these results are based on a short critical thinking task, improvement such as this could be replicated on a larger scale. It is likely that simply practising short tasks to improve critical thinking would improve students' performance on longer tasks and pieces of examined coursework such as the Critical Review. It may therefore be worthwhile for some departments to consider implementing a course of short critical thinking tasks such as the one used in this study, which could be of great benefit to their students.

As this study also found, carrying out the full reciprocal peer critiquing task significantly improved subjects' performance. With such significant differences between the control group and those carrying out the full task, even on this simple exercise, similar results are likely to be replicated if performance was to be studied when peer critiquing is used with the Critical Review or a similar piece of coursework. This is evidence that some tutors may need in order to convince them that using the reciprocal peer critiquing exercise is extremely advantageous to students. If more tutors employed the use of the task, students' coursework marks would improve, along with the other benefits gained in terms of self esteem etc (e.g. Bruffee, 1978; Falchikov, 1996).

Showing that reciprocal peer critiquing makes a significant improvement to students' work suggests it would be in the interest of students to employ such exercises as early in their university course as possible. By encouraging students in their first and second years to participate in tasks similar to this one, and letting them see the results they can gain, the students themselves should be more willing to organise peer critiquing amongst groups of friends, or within study groups, when they reach honours years. This could potentially result in a higher standard of work, and better results.

Future Directions

This study could be replicated on a larger scale in order to investigate if any one aspect of the peer critiquing task was found to make a significant improvement to average scores. By increasing the sample size, and by adjusting the questions to make them more suitable to the specific course, it is possible that significant results could be found for a single aspect, such as receiving feedback from peers. If a single aspect did prove to be significant it could make the exercise easier and more appealing to organise within a department, as well as being relatively quick for the students who participate in the task, making it ideal to fit into a one-hour tutorial, for example.

Organising a version of the reciprocal peer critiquing task for first and second year courses would take time, but with particular reference to the University of Glasgow Psychology Department, there are several pieces of coursework

within the first two years that would be ideal to trial peer critiquing on. By comparing results against previous years, any increase in standard of work could be noticed, and could be attributed to the peer critiquing exercise. Teaching students in first and second year methods of improving the standard of their work would have a positive effect on their work in future years, as well as promoting critical thinking.

Conclusion

This study has produced important findings in support of the use of reciprocal peer critiquing. The results showed that all groups showed some improvement after three attempts at the task, and subjects who participated in the full peer critiquing exercise showed significantly better results than controls after two intervention trials. Showing that a significant improvement can be made, even on a short task such as was employed in this study, will hopefully encourage more tutors in the Department of Psychology at the University of Glasgow to implement the reciprocal peer critiquing exercise with their students who are working on Critical Reviews. If significant support was found for use on the Critical Reviews, this could encourage tutors in other departments, and in other institutions, to implement reciprocal peer critiquing exercises.

This study did not find support for any single aspect of the exercise having a significant effect on average score on the critical thinking task. However, the small sample size and difficulties in attributing marks to relatively short passages could have affected the results, so further studies could be carried out to investigate these issues.

Acknowledgement

Thanks to Dr Steve Draper for his advice, comments and suggestions. His enthusiasm was encouraging, and his help was greatly appreciated.

References

Anderson, T., Howe, C., Soden, R., Halliday, J. & Low, J. (2001). Peer interaction and the learning of critical thinking skills in further education students. *Instructional Science, 29,* 1-32.

Assiter, A. (Ed) (1995). *Transferable Skills in Higher Education*. London, Kogan Page.

Boud, D., Cohen, R. & Sampson, J. (1999). Peer learning and assessment. *Assessment & Evaluation in Higher Education, 24 (4),* 413-426.

Bruffee, K.A. (1978). The Brooklyn Plan: attaining intellectual growth through peer group tutoring. *Liberal Education, 64, (4),* 447-468.

Davies, P. (2000) Computerised peer assessment *Innovations in Education and Teaching International* 37(4) 346-355

Draper, S.W. (2007). *Reciprocal Critiquing Exercise for Tutorials*. http://www.psy.gla.ac.uk/~steve/resources/tactics/tutcr.html [accessed 16/02/08].

Dunn, D.S. (1996) Collaborative writing in a statistics and research methods course *Teaching of Psy* 23 (1) 38-40

Falchikov, N. (1996). Improving learning through critical peer feedback and reflection. *Higher Education Research and Development, 19,* 214-218.

Falchikov, N. (2001). *Learning Together: Peer Tutoring in Higher Education*. London Routledge Falmer.

Hanrahan, S.J. & Isaacs, G. (2001). Assessing self- and peer-assessment: the students' views. *Higher Education Research & Development, 20 (1),* 53-70.

Morrow, L.I. (2006). An application of peer feedback to undergraduates' writing of critical literature reviews. *Practice and Evidence of Scholarship of Teaching and Learning in Higher Education, 1 (2),* 61-72.

Ozogul, G & Sullivan, H. (2007). Student performance and attitudes under formative evaluation by teacher, self and peer evaluators. *Education Technology Research and Development, doi:* 10.1007/s11423-007-9052-7

Topping, K.J., Smith, E.F., Swanson, I. & Elliot, A. (2000). Formative peer assessment of academic writing between postgraduate students. *Assessment Evaluation in Higher Education, 25 (2),* 149-167.

Venables, A. & Summit, R. (2003). Enhancing scientific essay writing using peer assessment. *Innovations in Education and Teaching International, 40 (3,),* 281-290.

Appendix 1: Passage Titles

Passage 1.

The legal drinking age should be lowered to 16 – discuss.

Passage 2.

Children nowadays are wrapped in cotton wool – discuss.

Passage 3.

The Olympics are a waste of money – discuss.

Appendix 2: Experiment Instructions

Experiment Instructions

This study has been designed in order to investigate the factors affecting critical thinking. I would like you to write 3 short passages in the form of mini critical reviews. By this, I mean I would like you to give a *balanced argument* on the topic. Each passage will be on a different topic, the title of which you will be provided with.

You will be allowed five minutes to complete each passage. There is a clock on display, and you will be informed when there is one minute, and then 30 seconds to go. Significant effort is required, despite the fact you are only allocated a short amount of time. Pilot studies showed that on average, participants managed to hand-write 14 lines of text.

When it is indicated that the five minutes are up, you will be told to turn the page.

Appendix 3: Structured Feedback Sheet

Structured Feedback Sheet

Student author (matric no.): Student assessor (matric no.):

1. Please comment on the author's inclusion of multiple views on the topic i.e. did they include more than one view?

2. Please comment on the reasons given by the author i.e. did they include reasons on both sides of the topic?

3. Please comment on the author's inclusion of a statement showing their clear support for **one side** of the topic.

4. Please comment on the author's piece of work as a whole, including writing style and overall quality of the work.

Appendix 4: Marking Scheme used for Passage Answers

Critical Thinking Marking Scheme

All pieces of work should be marked using the following marking scheme.
Allocate a suitable number of marks for each objective in the marking scheme, ranging from 0 (has not fulfilled the objective) to 5 (has fulfilled the objective very well).
Total the number of marks to give an overall score out of 20.

The objectives are as follows:

1. Inclusion of several points.
2. Inclusion of reasons on *both sides* of the topic.
3. Inclusion of a statement showing clear support for *one side* of the topic.
4. Text as a whole, taking into account writing style and quality of work.

Objective Number	**Mark**
1.	/5
2.	/5
3.	/5
4.	/5
Total	/20

Appendix 5: Participant Consent Form

Consent Form

This study is being carried out as part of the course requirements for the BSc Honours degree in Psychology at the University of Glasgow.
Participation in this study is voluntary, and you may withdraw your participation at any point without having to provide a reason.
You are assured that all raw data collected in this study will only be seen by the investigators, and all participants will remain anonymous in any published data.
By participating in this study, you are agreeing to write 3 short passages on given titles, and depending on the group you are allocated to, perhaps provide feedback on others' work.

If you consent to participating in this study please sign below:

Signature: Date:

Alison Gemmell | Supervised by Dr Steve Draper

Direct correspondence to 0404036g@student.gla.ac.uk
Alternatively to s.draper@psy.gla.ac.uk

Appendix 6: Participant Information Sheet

Participant Information Sheet

This study is being carried out as part of the course requirements for the BSc Honours degree in Psychology at the University of Glasgow.
Participation in this study is voluntary, and you may withdraw your participation at any point without having to provide a reason.

You are assured that all raw data collected in this study will only be seen by the investigators, and all participants will remain anonymous in any published data of the mean results.

The aim of this study is to look at different factors affecting students' performance on a short critical thinking task. You will be randomly assigned to one of five groups and asked to partake in a miniature critical thinking task, in which the conditions of practice and feedback will be varied depending on the group number. The task will be carried out in a quiet location using paper and pen. The group results will be compared to see where the biggest effect lies.

Please feel free to contact myself or my supervisor at any point if you have further questions or comments about the study.

Alison Gemmell | Supervised by Dr Steve Draper

Direct correspondence to 0404036g@student.gla.ac.uk
Alternatively to s.draper@psy.gla.ac.uk

Appendix 7: Table of Raw Data of each Subject's Scores

Group No.	Passage 1	Passage 2	Passage 3
1	9	9	9
1	11	12	12
1	11	10	11
1	12	11	13
1	9	12	13
2	13	14	18
2	9	4	9
2	14	13	15
2	7	11	14
2	8	10	11
3	11	13	14
3	13	14	12
3	9	14	13
3	12	15	16
3	10	14	15
4	9	14	14
4	10	12	15
4	7	10	13
4	11	8	9
4	9	10	12
5	8	10	13
5	11	12	16
5	13	16	16
5	9	13	15
5	10	14	17

Appendix 8: Statistics Output for Repeated Measures ANOVA using SPSS version 15

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Passage_no	Sphericity Assumed	128.000	2	64.000	29.268	.000
	Greenhouse-Geisser	128.000	1.496	85.563	29.268	.000
	Huynh-Feldt	128.000	1.913	66.908	29.268	.000
	Lower-bound	128.000	1.000	128.000	29.268	.000
passage_no * grp_no	Sphericity Assumed	34.533	8	4.317	1.974	.075
	Greenhouse-Geisser	34.533	5.984	5.771	1.974	.101
	Huynh-Feldt	34.533	7.652	4.513	1.974	.079
	Lower-bound	34.533	4.000	8.633	1.974	.137
Error(passage_no)	Sphericity Assumed	87.467	40	2.187		
	Greenhouse-Geisser	87.467	29.920	2.923		
	Huynh-Feldt	87.467	38.261	2.286		
	Lower-bound	87.467	20.000	4.373		

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	10443.000	1	10443.000	1025.164	.000
grp_no	66.267	4	16.567	1.626	.207
Error	203.733	20	10.187		

Appendix 9: Statistics Output for One-Way ANOVA using SPSS version 15

ANOVA

passage3

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	36.100	1	36.100	14.157	.006
Within Groups	20.400	8	2.550		
Total	56.500	9			