

## To give is better than to receive: The benefits of peer review to the reviewer's own writing

Kristi Lundstrom, Wendy Baker\*

*Department of Linguistics and English Language, Brigham Young University, Provo, UT, USA*

---

### Abstract

Although peer review has been shown to be beneficial in many writing classrooms, the benefits of peer review to the reviewer, or the student giving feedback, has not been thoroughly investigated in second-language writing research. The purpose of this study is to determine which is more beneficial to improving student writing: giving or receiving peer feedback. The study was conducted at an intensive English institute with ninety-one students in nine writing classes at two proficiency levels. The “givers” reviewed anonymous papers but received no peer feedback over the course of the semester, while the “receivers” received feedback but did not review other students’ writing. An analysis in the gains in writing ability measured from writing samples collected at the beginning and end of the semester indicated that the givers, who focused solely on reviewing peers’ writing, made more significant gains in their own writing over the course of the semester than did the receivers, who focused solely on how to use peer feedback. Results also indicated that givers at the lower proficiency level made more gains than those at higher proficiency levels and that slightly more gains were observed on global than local aspects of writing.

© 2008 Elsevier Inc. All rights reserved.

*Keywords:* Second-language writing; Peer review; Peer editing; Peer feedback

---

### Introduction

Peer review (i.e., peer editing, peer evaluation, or peer response), frequently used in both first (L1) and second language (L2) writing classrooms, is an important activity which allows writing teachers to help their students receive more feedback on their papers as well as give students practice with a range of skills important in the development of language and writing ability, such as meaningful interaction with peers, a greater exposure to ideas, and new perspectives on the writing process (Hansen & Liu, 2005; Mangelsdorf, 1992). Fundamental issues relating to peer review, such as how to train students (Hansen & Liu, 2005; Schaffer, 1996), how to form groups (Hansen & Liu, 2005; Mendonça & Johnson, 1994; Rollinson, 2005), the types of activities to conduct (e.g. Bell, 1991; McMurry, 2004; Schaffer, 1996), and the methods to be used (O’Donnell, 1980; Rollinson, 2005) are all dependent on the unique needs of the students involved. The many choices available to teachers when setting up peer review can be daunting, especially since what method is best varies with the situation. Thus, the adaptability of peer review can actually create confusion for teachers as to what exactly peer review involves and the best way to utilize it. Because of these

---

\* Corresponding author at: Brigham Young University, Department of Linguistics and English Language, 4057 JFSB, Provo, UT 84602, USA. Tel.: +1 801 422 4714; fax: +1 801 422 0906.

*E-mail address:* [kristilundstrom@gmail.com](mailto:kristilundstrom@gmail.com) (W. Baker).

challenges, some critics argue that peer review is of questionable value in the L2 classroom (e.g., Nelson & Murphy, 1992).

Many studies, however, support the idea that peer review can be extremely effective for a variety of reasons when used correctly (Bruffee, 1978; Lockhart & Ng, 1995; Paulus, 1999), especially when students are trained on how to give and use feedback (Min, 2006). Teachers can incorporate it as a way to present writing skills to students, ideally creating a student-centered classroom with learners capable of critically evaluating their own written work (Bell, 1991; Braine, 2003; Tang & Tithcott, 1999). Peer review for L2 learners also provides students with the opportunity to use language in the classroom in a meaningful way (Krashen, 1982), thus improving not only their writing but also allowing them to practice their listening and speaking abilities (Lockhart & Ng, 1995; Mendonça & Johnson, 1994; Tang & Tithcott, 1999). Peer review sessions can teach students important writing skills, such as writing to a real audience (Mangelsdorf, 1992), seeing ideas and points of view other than their own (Paulus, 1999), and discussing how to revise writing effectively (Lee, 1997). Finally, peer review teaches international students how to work in groups with their peers, a skill they may not have learned in their native country, but that is necessary for success in American universities and workplaces (Tang & Tithcott, 1999).

#### *Benefits of peer reviewing for the reviewer*

One aspect of peer review that could also provide extensive gains but that has rarely been investigated empirically in L2 writing research is the possible benefits of peer review to the giver, or the person reviewing the essay and offering feedback. The skill of being able to critically evaluate writing, defined as the ability to look at a classmate's writing and then provide effective feedback, particularly on a global level (i.e., at the level of content and organization), is a very necessary skill for quality writing and academic success in general (Gieve, 1998; Thompson, 2002). Developing critical evaluation skills may also help students effectively review texts and see logical gaps, problems with organization, and other defects that weaken the argument of the paper on a global level (Beach, 1989; Ferris, 2003; Thompson, 2002), making students better writers and self-reviewers.

Such skills may best be understood within the framework of sociocultural theory (e.g., Vygotsky, 1986), which theorizes that learners can only acquire information within their zone of proximal development (ZPD). The learner's ZPD refers to the place between where learners (i.e., "novices") are able to perform a task on their own versus with the help of a teacher or parent ("experts") (Aljaafreh & Lantolf, 1994). The goal is to help learners reach a level where they are able to perform a task on their own. Past studies cite joint scaffolding (i.e., each learner helping the other by providing extra support) to explain the effectiveness of peer review (i.e., De Guerrero & Villamil, 2000; Teo, 2006; Warwick & Maloch, 2003). For example, De Guerrero and Villamil (2000) demonstrate how two students, one the writer and one the reviewer, learn from each other during a peer review exercise. In this analysis, they demonstrate how at times the reviewer scaffolds the learning of the writer while at other times the writer scaffolds the learning of the reviewer. Therefore, one of the important findings of these studies is that even when two novice learners are paired together they still scaffold each other's learning (Anton & DiCamilla, 1998; Teo, 2006). Thus, both the giver and receiver of peer feedback may benefit from peer review activities.

Learning to effectively review others' writing may then ultimately lead to the creation of better self-reviewers, or students who are able to look at their own papers and accurately assess areas in which they need to improve and revise them (Rollinson, 2005). One study that suggests this may be the case was conducted by Min (2005), which found that the majority of the student participants commented that training on how to review their peers' papers helped them improve their own writing. One question that arises from this study, however, is whether or not the students' self-perceptions were accurate, and, if so, what specific aspects of the students' writing improved. Most studies that have been conducted, however, have not attempted to address this issue, but have instead focused on self-reported beliefs that peer review helped in one's own writing ability, such as in the Min (2005) study.

#### *L1 research on critical thinking and peer review*

Even in L1 research the question of whether or not peer review helps the reviewer does not seem to have been investigated in any known studies for almost fifteen years. Initially, discussion concerning the role of students in the classroom with respect to each other developed in the mid-1970s with the idea of collaborative learning. Bruffee

(1973) published a landmark article describing teachers as organizers of students rather than dispensers of knowledge. Later, other researchers built on Bruffee's ideas to argue that writing teachers should teach students how to edit papers collaboratively with their peers, with the assumption that this would allow them to be better editors of their own work (Butler, 1981; Gebhardt, 1980; O'Donnell, 1980).

In the same year that Bruffee published his article about collaborative learning, Sager (1973) published a study which further supports the idea that students who evaluate writing become better writers. She found that students who learned how to evaluate writing using a rubric showed significantly greater improvement than those who did not, and the students were 99 percent reliable in their assessments of the writing they evaluated. She concluded that teaching the students how to use the scale made them better judges of writing and also brought these issues to their attention more forcefully in their own writing, as measured by improvement in their own writing scores over the course of the study.

Questions began to arise, however, as to the truth of the claim that teaching students how to review their peers' papers actually teaches them to critically evaluate their own work and therefore become better writers (Matsubashi, Gillam, Conley, & Moss, 1989). Because tutors are in a unique situation in which they give but do not receive feedback, they offer an ideal test case into whether merely giving feedback affects one's own ability to write well. Two major studies looked at the writing of writing tutors in two schools. In particular, Bruffee (1978) examined the writing of tutors over a semester and found that both the tutors and the students they worked with improved significantly in their writing ability over the semester. Marcus (1984) conducted a similar study at a secondary school and found similar results. Graner (1987) also addressed this question in a study in which he compared two L1 writing classes to monitor the effects of peer review on the reviewer. He found that students who reviewed papers but did not receive any feedback on their own work improved at the same rate as students participating in traditional peer review activities, once again suggesting that, for writing students, giving feedback is at least, if not more, beneficial than receiving it.

#### *Need for L2 research*

Even though these studies provide a preliminary understanding of the benefits of reviewing L1 peer writing on a student's own L1 writing ability, there are still difficulties in assuming that these benefits would be similar in the L2 writing classroom, where language and culture may add unanticipated challenges (Carson & Nelson, 1994; Ramanathan & Atkinson, 1999). Specifically, the L1 studies cited above do not address complications introduced by differences in language proficiency or cultural expectations of students. For example, Kamimura (2006) found that while students at two different proficiency levels benefited from peer feedback activities, they differed in how they understood and used peer feedback. In addition, some students accustomed to a very teacher-fronted classroom may not feel comfortable working with peers in a more student-centered environment and may resist group-centered peer review activities, preventing them from developing the advanced critical evaluation skills thought to be associated with these activities (Braine, 2003). In addition, most of the studies do not use experimental research methods (i.e., they do not quantitatively compare two groups), nor do these studies examine the types of improvements students made, whether in global (i.e., organization, development, and cohesion) or local (i.e., grammar, vocabulary, and mechanics) aspects of writing. These complications need to be addressed before assuming that the findings of L1 studies can be generalized to L2.

The need to understand how writing teachers can maximize the benefits of peer review, including to what extent Bruffee's (1973) findings can be generalized to the L2 writing classroom, led to the current study. Specifically, this study will address the following questions:

1. Do students who review peer papers improve their writing ability more than those who revise peer papers (for both beginning and intermediate students)?
2. If students who review peer papers do improve their writing ability more than those who revise them, on which writing aspects (both global and local) do they improve?

To our knowledge, no rigorous empirical studies have been done in L2 research to show that the act of reviewing peer written work really does improve students' ability to critically evaluate writing, an ability which is then transferred into the students' own writing process, resulting in better writing on both local and global levels. This leaves a gap in the research on peer review that would benefit L2 writing teachers struggling to provide sufficient quality feedback to their students, because this research could allow teachers to know that students not only benefit each other through peer feedback but also benefit themselves.

## Methodology

### *Participants*

The participants in this study were ninety-one students enrolled in nine sections of writing classes at the English Language Center (ELC) at Brigham Young University. At the ELC, students progress through five levels with level one corresponding to a beginning level class and level five to a low advanced class. Forty-five of the participants in the study were enrolled in level 2 (high beginning) and forty-six in level 4 (high intermediate) writing classes. At both proficiency levels, the participants were divided into two groups. The first group was composed of two high beginning and three high intermediate classes, (totaling forty-six students). This group was the control group (hereafter “receivers”) and received peer feedback but did not review peers’ papers (defined as compositions written by students at their same proficiency level). The second group was composed of two high beginning classes and two high intermediate classes, (totaling forty-five students), and made up the experimental group, who reviewed peer papers but did not receive peer feedback (hereafter “givers”). Class sizes ranged from twelve to seventeen. The students spoke eight different native languages, and 46% were male and 54% were female.

### *Procedure*

As discussed above, the participants were divided into the control group (“receivers”) and the experimental group (“givers”), based on which class they were in. The givers reviewed peer papers but not receive peer feedback, while the receivers received peer feedback but did not give it. The experimental group, therefore, was similar to that of the L1 writing studies cited above, where writing tutors gave feedback but did not provide it (Bruffee, 1978; Marcus, 1984). Unlike earlier studies, we also included a group that revised papers but did not review peer papers. By having the two groups receive these different treatments, we were able to address the question of where the benefits of peer review can be found, whether in the giving or the receiving of feedback or both, in order to help writing teachers use peer review activities more effectively.

The procedure was as follows: Four times throughout the semester, participants received training on peer review. The only difference between the giver and receiver training was that the exercises that the receivers were given focused on how to use feedback to revise a paper, whereas the givers were given instruction on how to give feedback. For example, both groups received a lesson on how to write effective introductions, which is a regular part of the writing curriculum. In addition, the receiver group was taught how to use feedback to revise an introduction and the giver group how to give effective feedback on introductions. Next both groups were given the same example essay written by a student at their proficiency level. Along with this essay, both groups were given the same list of questions. For instance, in the example above, one of the corresponding questions was “How would you improve the thesis statement?” The receivers were to take the feedback written in the margins of the example essay and then rewrite the thesis statement. The “givers” were to provide their own suggestions for improving the thesis statement and write it in the margins of the example essay (their copy of the essay did not include suggestions already written in the margins). Thus, both groups were asked to look at the same issues in the essay, but the receivers were to revise the paper and the givers were asked to provide suggestions on how to improve the paper.

Importantly, although students did write compositions in their classes, the only peer review they participated in was these sample essay exercises. These essays were authentic student texts written within the last year. We used these texts instead of having students use their own papers in order to control for differences in student writing (since with different papers there would be the possibility of wide differences in both how well the papers were written and what types of changes were needed). We also did not have students peer review each other’s papers because we wanted to ensure that the giver group only gave feedback and the receiver group only learned how to use (but did not give) peer feedback.

Using the same essays and lesson plans for all classes helped ensure that the students were receiving similar instructions. These lesson plans supported the writing objectives at the English language institute and could be adapted by the teachers at their discretion, but teachers could not change the plans significantly. For example, in one of the lesson plans, the teachers were asked to explain how to write a paragraph using main ideas and details. The teachers were also asked to explain how to write an effective paragraph, model writing a paragraph with the class, and then have each student write a paragraph on his or her own. The teachers could change the topic used for the model essays, but

they could not necessarily change what was taught about how to receive or give feedback. This allowed for some adaptability, although each class had to teach the basic points of the lesson. After both groups finished the sample essay assignment, teachers were not told any specific differences in dealing with the assignments.

When two treatments, as in this study, are given to students at the same institution, it is possible that the two groups conversed with each other and that one group may have been unhappy that they received a different treatment from the other. To our knowledge, such complaints were minimal if any—especially since (as described above) most of the instruction to the two groups was similar. The only differences occurred in the application of the lessons given.

### *Raters*

As pre- and post-tests, students were asked to write a timed essay. These essays were rated by seven teachers (two males, five females) working at the English language institute who were either current or past writing teachers with one to eight semesters of previous experience teaching writing. All had experience in grading student essays as well as rating final writing portfolios using the institution's five-point rubric. Five raters scored the student essays from the first semester of data collection, and five raters scored the essays from the second semester, including several raters scoring essays both semesters. It would have been ideal to have the same raters for all the essays, but this was not feasible given the raters' teaching schedules and time constraints. Therefore, we tried to ensure rater reliability by having two raters who scored all the essays. In addition, as described below several analyses were performed to ensure that the ratings were calibrated and the instrument was valid. All raters rated essays written by both beginning and intermediate students, and the distribution of essays to raters was completely random.

### *Rating rubric*

To assess the writing proficiency of the students, the grading rubric used by Paulus (1999) was used in this study (see Appendix A). The scale was used to assess the thirty-minute diagnostic essay (pre-test) and the thirty-minute final essay (post-test). This scoring guide was chosen because it allowed for an analytical assessment of both the global and local aspects of writing, in addition to providing a holistic, overall final assessment score. The scoring guide is based on a ten-point scale, and student essays were assigned a score on the scale from 1 to 10 for each of the following writing aspects: organization, development, cohesion, structure, vocabulary, and mechanics. In this scale, organization refers to the effectiveness of the thesis statement and unity of ideas, development refers to the appropriate use of examples and support, cohesion refers to the relationship of ideas to each other and the use of transitions, structure refers to syntax complexity and grammatical accuracy, vocabulary refers to clarity of meaning and the precision of the words used, and mechanics refers to spelling, punctuation, capitalization, and general formatting. The criteria on the rubric were then averaged to assign an overall score.

We attempted to ensure the validity of this rubric in three ways. The first was to have raters practice using the rating rubric on five thirty-minute diagnostic or final essays similar to those used in the study but written by students from previous semesters. These essays were rated by each rater, and then the raters compared their scores and discussed their reasons for choosing their rating until all the raters agreed on a given score. They repeated this process until all the raters consistently rated within one point on a ten point scale, understood the six criteria measured by the scale, and felt comfortable using the whole range of the scale.

Second, all the essays were graded by at least two raters and the different scores were then averaged. If the first two raters disagreed by more than one point in any of the seven criteria of a given essay, a third rater also graded the disputed aspects of the essay before the scores for the essay were averaged. To more accurately determine the reliability of the raters, each semester several essays were scored by all the raters.

Finally, to ensure that all the scores were valid we performed a FACETS analysis and used the FACETS scores to run all statistics. This procedure for data analysis is explained below.

### **Data analysis**

To examine whether the givers improved their writing more than the receivers throughout the semester, each student wrote a thirty-minute timed essay collected the first week of class and a final thirty-minute essay collected the last week of the semester. These essays served as pre- and post-test assessments. In-class timed essays were used as a

measure of the students' writing ability to ensure that the students did their own work and did not receive any outside help. The topics of the diagnostic essays were assigned by the writing coordinator according to the writing objectives for that level, and they were the same for all classes of that proficiency level.

The data collected from rating the timed essays were sorted by proficiency level (beginning and intermediate) and group (givers or receivers). The data included an overall score for each pre- and post-test essay, as well as the scores for the 6 aspects of writing: organization, development, cohesion, structure, vocabulary, and mechanics. The data were next run through FACETS, a statistical program based on the Multi-Faceted Rasch Analysis (MFRA) model, which is in turn founded in item response theory (Lynch & McNamara, 1998; Park, 2004; Winsteps, 2008). By using FACETS, the scores that were used to calculate gain (i.e., the difference between the pre- and post-test scores), called logit scores, were adjusted for rater severity or leniency and were more accurate representations of student writing ability independent of raters. The results of the FACETS analysis also indicated that the students' essays were scored consistently. Once the logit scores were obtained for all the students' diagnostic and final essays, the statistical analyses were run using these scores. Therefore, this analysis determined the validity of the grading method.

## Results

The first question of this study was whether the giver group, which focused on reviewing peer essays, made larger gains in their writing scores from pre- test to post-test than did the receiver group, which focused on using peer feedback. To answer this question, we performed a repeated measures ANOVA with the 7 writing aspects (overall, organization, development, cohesion, vocabulary, mechanics, grammar) and time (pre-test and post-test) as within subject matters and treatment group (receiver vs. giver) as between subjects factors. These analyses determined which of the two groups made significant gains from pre- test to post-test, and they were run separately from for the beginning and intermediate levels to determine if differences occurred on both levels. The second question of this research study asked: if differences exist between the two groups, on which aspects of writing do they occur (global or local aspects)? Follow-up post-hoc Tukey tests were run to answer this research question.

### *Beginning learners*

To examine these two questions for the beginning groups, a repeated measures ANOVA was run on the beginning giver and receiver groups' pre-test and post-test scores with each writing aspect (overall, organization, development, cohesion, vocabulary, structure, and mechanics) and time (pre- vs. post-test) as within and treatment (giver vs. receiver) as between subject factors. The results of this analysis revealed a significant effect of time ( $F(1,45) = 238.805, p < .0001$ ), a time  $\times$  aspect ( $F(1,6) = 8.27, p < .001$ ), and a treatment  $\times$  time interaction ( $F(1,6) = 28.209, p < .0001$ ), suggesting that both groups improved in some aspects of the writing process, but that one group had greater gains than the other group on some of these aspects. Further post-hoc analyses indicated that the beginning giver group had greater significant gains for overall and all three global aspects than did the beginning receiver group.

The results of the repeated measures ANOVA suggested that both groups had gains in at least some of their writing skills from pre- test to post-test, but that the beginning group had greater gains than did the receiver group. One concern about these results is that the beginning giver group's pre-test scores appear to be lower than those for the receiver group (see Table 1). A two-way (treatment  $\times$  writing aspect) ANOVA run on the pre-test scores of the receiver versus the giver groups revealed a main effect of treatment ( $F(1,45) = 9.014, p < .003$ ), suggesting that the receiver group had better scores on the pre-test than the giver group. This would be problematic if the giver group did not also have higher post-test scores than the receiver group, since it would suggest that the differences in gains were the result of differences in pre-test scores only. However, a similar analysis was run on the post-test scores which revealed that the giver group had higher post-test scores than the receiver group (significant effect of treatment:  $F(1,45) = 15.899, p < .0001$ ), and that this occurred for every writing aspect, since there was no writing aspect  $\times$  treatment interaction ( $F(1,45) = .101, p > .05$ ). In other words, the giver group not only started lower in testing scores than did the receiver group, but they also outperformed the receiver group on all post-test scores (Table 1).

Such findings suggest that, at least for the beginning group, giving peer feedback on a paper seems to improve writing ability more than learning to use peer feedback to improve writing. Moreover, this analysis revealed that the beginning giver group outperformed the receiver group in all global writing aspects.

Table 1

Pre- and post-tests (and the difference between them) for the beginning givers and receivers groups (FACETS logit pre- and post-test scores given below each raw score; standard deviations given in parentheses)

	Receivers (C)			Givers (T)		
	Pre-test	Post-test	Diff.	Pre-test	Post-test	Diff.
Organization	3.39 (.80)	4.64 (1.03)	1.25	<b>3.08 (.88)</b>	<b>4.89 (.70)</b>	<b>1.81</b>
	–2.37 (2.73)	0.41 (2.75)	2.78	<b>–3.03 (2.87)</b>	<b>1.31 (1.86)</b>	<b>4.34</b>
Development	3.20 (.91)	4.26 (1.11)	1.06	<b>3.00 (1.00)</b>	<b>4.60 (.84)</b>	<b>1.60</b>
	–1.88 (2.67)	–0.01 (2.33)	–2.39	<b>–2.39 (3.13)</b>	<b>0.86 (1.75)</b>	<b>3.25</b>
Cohesion/Coherence	3.05 (.65)	3.90 (.91)	0.85	<b>2.85 (.92)</b>	<b>4.39 (1.01)</b>	<b>1.54</b>
	–1.55 (1.92)	–0.29 (1.72)	1.26	<b>–2.34 (2.97)</b>	<b>0.49 (1.51)</b>	<b>2.83</b>
Structure	2.92 (.78)	3.71 (.66)	0.79	2.54 (.71)	4.03 (1.05)	1.49
	–2.54 (2.27)	–1.00 (1.24)	1.54	–3.32 (2.17)	–0.53 (1.65)	2.79
Vocabulary	3.19 (.67)	4.01 (.75)	0.82	2.75 (.72)	4.25 (.94)	1.50
	–1.36 (1.76)	–0.53 (1.49)	0.83	–2.99 (2.77)	0.23 (1.99)	3.22
Mechanics	3.20 (.74)	3.75 (.89)	0.55	2.93 (.90)	4.15 (.93)	1.22
	–1.82 (1.83)	–0.83 (1.73)	–2.99	–2.41 (2.38)	–0.04 (1.65)	2.37
Overall	3.16 (.74)	4.04 (.91)	0.88	<b>2.91 (.77)</b>	<b>4.38 (.94)</b>	<b>1.47</b>
	–1.27 (1.15)	–.45 (1.03)	0.82	<b>–1.76 (1.51)</b>	<b>0.09 (1.03)</b>	<b>1.85</b>

\* $p < .001$  (significant differences in bold).

### Intermediate learners

A similar repeated measures ANOVA was run on the intermediate learners' pre-test and post-test scores with each writing aspect (overall, organization, development, cohesion, vocabulary, structure, and mechanics) and time (pre- vs. post-test) as within and treatment (giver vs. receiver) as between subject factors. The results of this analysis revealed a significant effect of time ( $F(1,43) = 7.926, p = .005$ ) but no effect of treatment ( $F(1,43) = .370, p > .05$ ) nor a treatment  $\times$  time interaction ( $F(1,43) = 2.23, p > .05$ ), suggesting that both groups did improve from pre-test to post-test on some of the writing aspects, but that they did not differ from each other in gains. Further post-hoc analyses revealed that both the giver and the receiver intermediate group improved in organization, development, and structure but not in the other aspects (Table 2).

It is possible that there were no differences between the giver and receiver groups in the intermediate learners because they had spent several semesters learning and using peer review. By contrast, none of the beginning level

Table 2

Pre-test and post-test scores (and the difference between them) for the intermediate givers and receivers groups (FACETS logit pre- and post-test scores given below each raw score; standard deviations given in parentheses)

	Receivers (C)			Givers (T)		
	Pre-test	Post-test	Diff.	Pre-test	Post-test	Diff.
Organization	<b>4.91 (.90)</b>	<b>5.32 (1.15)</b>	<b>0.41</b>	<b>4.74 (1.15)</b>	<b>5.29 (1.38)</b>	<b>0.55</b>
	<b>2.27 (2.48)</b>	<b>2.32 (2.97)</b>	<b>0.05</b>	<b>1.61 (3.03)</b>	<b>1.61 (3.65)</b>	<b>0</b>
Development	<b>4.76 (.98)</b>	<b>5.15 (1.01)</b>	<b>0.39</b>	<b>4.80 (1.00)</b>	<b>5.04 (1.32)</b>	<b>0.24</b>
	<b>1.96 (2.05)</b>	<b>1.89 (1.83)</b>	<b>–0.07</b>	<b>2.03 (1.86)</b>	<b>1.44 (2.61)</b>	<b>–0.59</b>
Cohesion/coherence	4.62 (1.22)	4.72 (1.17)	0.10	4.60 (1.13)	4.80 (1.36)	0.20
	1.65 (2.10)	1.05 (1.86)	–0.6	1.62 (1.82)	0.91 (2.46)	–1.52
Structure	<b>4.73 (.91)</b>	<b>5.00 (1.14)</b>	<b>0.27</b>	<b>4.92 (1.05)</b>	<b>4.71 (1.14)</b>	<b>–0.21</b>
	<b>1.36 (1.78)</b>	<b>1.56 (1.82)</b>	<b>0.2</b>	<b>1.61 (1.83)</b>	<b>0.70 (2.12)</b>	<b>–0.91</b>
Vocabulary	4.80 (1.08)	4.97 (1.05)	0.17	4.87 (.86)	4.91 (1.08)	0.04
	2.09 (2.23)	1.92 (1.81)	–0.17	2.14 (1.70)	1.55 (2.09)	–0.59
Mechanics	4.69 (1.07)	4.61 (1.03)	–0.08	4.94 (1.01)	4.73 (1.11)	–0.19
	1.34 (2.32)	0.92 (1.91)	–0.42	1.72 (1.84)	0.95 (1.95)	–0.77
Overall	4.78 (.90)	4.97 (1.08)	0.19	4.84	5.09	0.15
	1.18 (1.23)	0.91 (1.16)	–1.08	1.21 (1.13)	0.67 (1.51)	0.54

\* $p < .007$  (significant items in bold).

learners had received previous peer review training. To determine whether exposure to earlier peer review training may have influenced our results, we divided the participants in both the treatment and control groups into new and returning learners. We hypothesized that those students for whom this was their first semester in an English class in the United States may have never been exposed to peer review. If this was the case, then perhaps the difference in gains between these two groups (new students in the givers versus new students in the receiver group) may show similar results to those found when comparing the beginning-level giver and receiver groups.

To test whether this was indeed the case, a similar analysis was performed on the intermediate giver and receiver participants' pre-test and post-test scores for whom this was their first semester in a writing class in the United States. A two-way (treatment  $\times$  writing aspects) ANOVA was run on the gains from pre-test to post-test. This analysis revealed a main effect of treatment ( $F(1,15) = 4.43, p < .05$ ), suggesting that for those students who were new to peer review, the intermediate giver students made greater gains than did the students in the receiver group. Post-hoc analyses revealed this was the case for the overall, organization, and development aspects. When the same analysis was run on the returning giver and receiver groups' gain scores, no significant differences were found ( $F(1,31) = .999, p > .05$ ). The results for the overall gain scores for the new and returning students in both the giver and receiver groups are shown in Fig. 1.

These above analyses addressed the impact that exposure to reviewing or revising papers had on whether students' writing improved significantly from pre-test to post-test. It also examined the specific areas in which students made significant improvements and whether those areas were largely global writing aspects or local aspects. The results of this analysis showed that beginning students had more significant gains in improvement than intermediate students. The beginning givers improved significantly in all areas, and they improved more than the receiver group in overall and the global writing aspects. By contrast, the intermediate givers and receivers made significant gains in fewer areas (organization, development, and grammar) and did not differ from each other in language gains on any of the writing aspects. However, when the intermediate students who had not participated in previous peer review at the ELC were separated from the returning students, it was found that the giver and receiver groups did differ significantly from each other in gains scores for overall and two of the three global writing aspects (development and organization).

It is possible that the gains found for the giver group over the receiver group were the result of factors other than differences in treatment, such as quality of teaching, individual student differences or different experience from pre-test to post-test. Although it is impossible to disregard these factors entirely, at least three points suggest that differences in treatment across the two groups did contribute in differences in gain scores for the giver versus receiver groups. First, we tried to control for teaching effects and individual differences in students by having eight different teachers and their students participate in this study. Our findings indicate that students in all the giver group classes made significantly greater gains than those in the receiver group classes. Second, we saw the same kinds of gains in the beginning group students as in the intermediate giver group who had no previous peer review experience (global areas). Moreover, these improvements were in the areas in which instruction differed between the giver and the receiver groups. If other factors were at play in differences across groups, we would not expect all the giver groups to make the same types of gains. Finally, it would be surprising for the intermediate students in the giver group who had never been exposed to peer review to differ so much in gains from similar students in the receiver group if it were not at least in part due to differences in treatment. Thus, while further research is needed to validate these findings, it appears

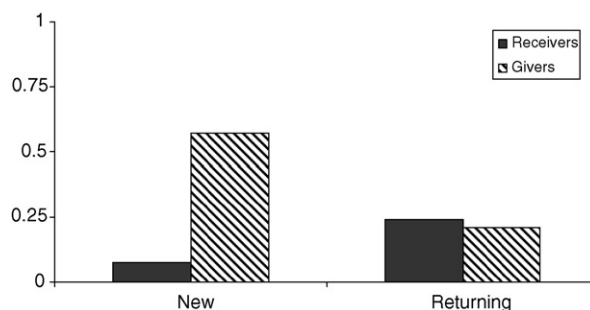


Fig. 1. Gain scores for the new intermediate givers and receivers (i.e., those who had not had peer review training in the ELC), versus gains for the returning givers and receivers (i.e., those who had previous experience with peer review training in the ELC).



that differences across the treatment groups were due at least in part to differences in whether or not the students gave or received peer feedback.

## General discussion

The findings of this study suggest that L2 writing students can improve their own writing by transferring abilities they learn when reviewing peer texts. In addition, these findings also suggest that students taught to give peer feedback improve in their own writing abilities more than students taught to use peer feedback. These results seem to support earlier findings in both L1 (e.g., Sager, 1973) and L2 (e.g., Teo, 2006) writing research. One reason that learning to review others' writing improves one's own writing may be that students learn from these activities to critically self-evaluate their own writing in order to make appropriate revisions (Rollinson, 2005). Such skills may not be developed when students are only taught how to interpret others' feedback, as was the case for the group in this study who only learned to use peer feedback.

The results of this study also may be explained in light of sociocultural theory discussed above. It is possible that reviewers of peer papers improve more than the revisers because the reviewers likely determine the level at which the peer review occurs. Specifically, reviewers often determine what aspects of writing that the peer review will focus on and most likely provide instruction for the writer that falls within their (the reviewer's) ZPD. If the writer's own ZPD is not at the level of the reviewer, he or she may not get feedback that scaffolds learning and therefore may not benefit as much from the review (Nassaji & Swain, 2000). Thus reviewers may be able to learn more from the feedback they give than writers can learn from the feedback they receive.

Of course in earlier studies examining scaffolding and peer review (i.e., De Guerrero & Villamil, 2000), one benefit of peer review was the interaction between the two students. In the present study no such interaction occurred in order to separate the giving and receiving of feedback. Certainly, future research should examine how this talk and interaction helps students who give peer review improve their writing and most likely would demonstrate even greater benefits for these students than occurred in this study. Nevertheless, it still holds that the students in the giver group in this study may have been better able to control what they learned within their ZPD than those in the receiver groups. This difference may explain why these two groups differed in how much they improved in writing aspects across the two semesters.

Such findings are significant since they suggest that peer review may be even more beneficial than thought before. It is not just the added feedback students receive on their writing, nor the extra language interaction experience that helps improve student writing; the act of providing feedback may also improve student writing and may be the most beneficial aspect of peer review. Finding that these benefits of peer review improve L2 writing matches earlier research that demonstrates that learning self-evaluation skills is beneficial in other aspects of L2 learning (i.e., Swain, Brooks, & Tocalli-Beller, 2002; Sullivan & Lindgren, 2002).

Results from this study indicated that those students who revised student papers improved in specific areas of writing more so than those who only learned to use student feedback. Since our main focus in this study was on whether or not students improved in global aspects of writing, we predicted that the givers would improve more on these aspects than the receivers. In actuality, we found this to be the case at least at the beginning level: the beginning giver and receiver groups differed in their improvement in both overall and global writing ability. In global writing aspects, the beginning givers improved significantly more than the receivers in all global writing aspects (organization, development, and cohesion). In addition, for those intermediate students who had had little experience in peer review at the ELC, the giver students made greater gains than the receiver students but only in overall as well as the global aspects of development and organization.

These findings may suggest that reviewing peer writing helps students learn global aspects of writing more than does learning how to interpret peer feedback. However, the lesson plans given to both groups focused solely on global aspects of writing, so this result is not surprising. However, further research that examines how peer reviewing versus revising affects local aspects of writing may illuminate whether similar benefits are found for writing aspects such as structure, vocabulary, and mechanics.

### *Proficiency level*

We also examined whether proficiency level affected the degree to which reviewing peer papers improved one's own writing ability. From these analyses, it was found that most of the differences between the givers and receivers

were obtained at the lower (beginning) giver group, which made greater significant gains than the receivers. Neither the intermediate givers nor receivers improved significantly from each other in any areas, but they did both improve in organization, cohesion, and vocabulary from pre-test to post-test. The beginning students' greater gain in overall writing ability could be interpreted to mean that students at intermediate proficiency levels do not benefit as much from peer review as do beginning students.

Several reasons may account for this difference in significant writing improvement between the beginning and intermediate students. One possible reason is that students at the beginning level, because their language skills were less developed, had more room for improvement than did the higher proficiency group, and therefore, the effects of new skills they developed resulted in a greater relative improvement in their writing ability. It is possible that this is also why the beginning giver group had higher gains than did the receiver group, although it does not explain why they had higher overall post-test scores than did the beginning receiver group. Also, students at the higher proficiency level may have had more experience with peer review than those at the lower level, pushing them further along the learning curve. This could mean that the more advanced students had already received the greatest initial benefits to be gained from peer review, and therefore, the improvement they made seemed less significant relative to where they started. In addition, the intermediate group may be developing skills that take longer than one semester to develop; thus the benefits from learning how to better revise student papers at their level will not be manifest over the course of only one semester. This later explanation seems plausible, since a further analysis of those intermediate students who had not had previous peer review training revealed that students in the giver group outperformed the students in the receiver group. Past research also suggests that peer review is more beneficial for students at one level than another (Kamimura, 2006).

Another explanation for these differences may be that beginning learners are still learning how to negotiate with their peer reviewer to receive feedback within their (the reviser's) ZPD, and therefore, they benefit more from reviewing than from receiving peer feedback. Peer review may be a process that requires writers to understand how to negotiate for instruction that falls within their ZPD. Indeed, Nassaji and Swain (2000) found that writers who received teacher feedback only within their ZPD improved more in writing ability than did writers who received random feedback, some of which was in their ZPD and some of which was not. Therefore, more advanced students may have learned to focus only on revision aspects that fall within their ZPD.

The findings of this study support the idea that reviewing other students' papers is a viable and important activity to improving one's own writing, findings which can benefit students on several levels (Bell, 1991; Paulus, 1999). By participating in these activities, students may develop the ability to critically examine even their own writing, which offers them self-feedback and greatly improves their writing skills. It is important, however, to remember that in this study the receivers did not receive peer feedback on their own writing nor did the givers interact with peers to give feedback. This was done so that the two tasks, giving and receiving feedback, could be separated enough to see their individual effects on students' writings. By doing this, the conditions of this study were not similar to a typical peer review situation. Thus, to greater support the findings of this study, further research should examine the effects of these two tasks in a qualitative study that closely identifies which aspects students discuss while in the reviewing part of peer review and whether these same aspects are improved in the reviewer's own writing. Such a study would complement the quantitative analysis provided in this study and would further illuminate how reviewing papers improves one's own writing.

In this same vein, it should be noted that the students in the receiver group also improved in both overall and in specific areas over the course of the semester. This suggests, as previous research has shown (e.g., Bruffee, 1978; Lockhart & Ng, 1995; Paulus, 1999; Hyland, 2000), that revising is also a beneficial activity and when combined with reviewing in a peer review session, students may gain even greater benefits than those found in the groups separately in this study. Thus, in a typical peer review session, students would ideally develop the thinking skills necessary to effectively evaluate a paper, as well as practice using feedback they receive from their peers. Therefore, although effective peer review activities take time and training to make them work, they can be very effective in developing student writers, particularly at lower proficiency levels or with those who have had little experience with peer review writing.

## Appendix A

### A.1. Essay-scoring rubric

	Organization/unity	Development	Cohesion/coherence	Structure	Vocabulary	Mechanics
1	No organization evident; ideas random, related to each other but not to task; no paragraphing; no thesis; no unity	No development	Not coherent; no relationship of ideas evident	Attempted simple sentences; serious, recurring, unsystematic grammatical errors obliterate meaning; non-English patterns predominate	Meaning obliterated; extremely limited range; incorrect/unsystematic inflectional, derivational morpheme use; little to no knowledge of appropriate word use regarding meaning and syntax	Little or no command of spelling, punctuation, paragraphing, capitalization
2	Suggestion of organization; no clear thesis; ideas listed or numbered, often not in sentence form; no paragraphing/grouping; no unity	Development severely limited; examples random, if given.	Not coherent; ideas random/unconnected; attempt at transitions may be present, but ineffective; few or unclear referential ties; reader is lost.	Uses simple sentences; some attempts at various verb tenses; serious unsystematic errors, occasional clarity; possibly uses coordination; meaning often obliterated; unsuccessful attempts at embedding may be evident	Meaning severely inhibited; very limited range; relies on repetition of common words; inflectional/derivational morphemes incorrect, unsystematic; very limited command of common words; seldom idiomatic; reader greatly distracted	Some evidence of command of basic mechanical features; error-ridden and unsystematic
3	Some organization; relationship between ideas not evident; attempted thesis, but unclear; no paragraphing/grouping; no hierarchy of ideas; suggestion of unity of ideas	Lacks content at abstract and concrete levels; few examples	Partially coherent; attempt at relationship, relevancy and progression of some ideas, but inconsistent or ineffective; limited use of transitions; relationship within and between ideas unclear/non-existent; may occasionally use appropriate simple referential ties such as coordinating conjunctions	Meaning not impeded by use of simple sentences, despite errors; attempts at complicated sentences inhibit meaning; possibly uses coordination successfully; embedding may be evident; non-English patterns evident; non-parallel and inconsistent structures	Meaning inhibited; limited range; some patterns of errors may be evident; limited command of usage; much repetition; reader distracted at times	Evidence of developing command of basic mechanical features; frequent, unsystematic errors
4	Organization present; ideas show grouping; may have general thesis, though not for persuasion; beginning of hierarchy of ideas; lacks overall persuasive focus and unity	Underdeveloped; lacks concreteness; examples may be inappropriate, too general; may use main points as support for each other	Partially coherent, main purpose somewhat clear to reader; relationship, relevancy, and progression of ideas may be apparent; may begin to use logical connectors between/within ideas/paragraphs effectively; relationship between/within ideas not evident; personal pronoun references exist, may be clear, but lacks command of demonstrative pronouns and other referential ties; repetition of key vocabulary not used successfully	Relies on simple structures; limited command of morpho-syntactic system; attempts at embedding may be evident in simple structures without consistent success; non-English patterns evident	Meaning inhibited by somewhat limited range and variety; often uses inappropriately informal lexical items; systematic errors in morpheme usage; somewhat limited command of word usage; occasionally idiomatic; frequent use of circumlocution; reader distracted	May have paragraph format; some systematic errors in spelling, capitalization, basic punctuation

5	Possible attempted introduction, body, conclusion; obvious, general thesis with some attempt to follow it; ideas grouped appropriately; some persuasive focus, unclear at times; hierarchy of ideas may exist, without reflecting importance; some unity	Underdeveloped; some sections may have concreteness; some may be supported while others are not; some examples may be appropriate supporting evidence for a persuasive essay, others may be logical fallacies, unsupported generalizations	Partially coherent; shows attempt to relate ideas, still ineffective at times; some effective use of logical connectors between/within groups of ideas/paragraphs; command of personal pronoun reference; partial command of demonstratives, deictics, determiners	Systematic consistent grammatical errors; some successful attempts at complex structures, but limited variety; clause construction occasionally successful, meaning occasionally disrupted by use of complex or non-English patterns; some non-parallel, inconsistent structures	Meaning occasionally inhibited; some range and variety; morpheme usage generally under control; command awkward or uneven; sometimes informal, unidiomatic, distracting; some use of circumlocution	Paragraph format evident; basic punctuation, simple spelling, capitalization, formatting under control; systematic errors
6	Clear introduction, body, conclusion; beginning control over essay format, focused topic sentences; narrowed thesis approaching position statement; some supporting evidence, yet ineffective at times; hierarchy of ideas present without always reflecting idea importance; may digress from topic Essay format under control; appropriate paragraphing and topic sentences; hierarchy of ideas present; main points include persuasive evidence; position statement/thesis narrowed and directs essay; may occasionally digress from topic; basically unified; follows standard persuasive organizational patterns	Partially underdeveloped, concreteness present, but inconsistent; logic flaws may be evident; some supporting proof and evidence used to develop thesis; some sections still undersupported and generalized; repetitive	Basically coherent in purpose and focus; mostly effective use of logical connectors, used to progress ideas; pronoun references mostly clear; referential/anaphoric reference may be present; command of demonstratives; beginning appropriate use of transitions	Some variety of complex structures evident, limited pattern of error; meaning usually clear; clause construction and placement somewhat under control; finer distinction in morpho-syntactic system evident; non-English patterns may occasionally inhibit meaning	Meaning seldom inhibited; adequate range, variety; appropriately academic, formal in lexical choices; successfully avoids the first person; infrequent errors in morpheme usage; beginning to use some idiomatic expressions successfully; general command of usage; rarely distracting	Basic mechanics under control; sometimes successful attempts at sophistication, such as semi-colons, colons
		Acceptable level of development; concreteness present and somewhat consistent; logic evident, makes sense, mostly adequate supporting proof; may be repetitive	Mostly coherent in persuasive focus and purpose, progression of ideas facilitates reader understanding; successful attempts to use logical connectors, lexical repetition, synonyms, collocation; cohesive devices may still be inconsistent/ ineffective at times; may show creativity; possibly still some irrelevancy	Meaning generally clear; increasing distinctions in morpho-syntactic system; sentence variety evident; frequent successful attempts at complex structures; non-English patterns do not inhibit meaning; parallel and consistent structures used	Meaning not inhibited; adequate range, variety; basically idiomatic; infrequent errors in usage; some attention to style; mistakes rarely distracting; little use of circumlocution	Occasional mistakes in basic mechanics; increasingly successful attempts at sophisticated punctuation; may have systematic spelling errors
8	Definite control of organization; may show some creativity; may attempt implied thesis; content clearly relevant, convincing; unified; sophisticated; uses organizational control to further express ideas; conclusion may serve specific function  Highly effective organizational pattern for convincing, persuasive essay; unified with clear position statement; content relevant and effective	Each point clearly developed with a variety of convincing types of supporting evidence; ideas supported effectively; may show originality in presentation of support; clear logical and persuasive/convincing progression of ideas Well-developed with concrete, logical, appropriate supporting examples, evidence and details; highly effective/convincing; possibly creative use of support	Coherent; clear persuasive purpose and focus; ideas relevant to topic; consistency and sophistication in use of transitions/referential ties; effective use of lexical repetition, derivations, synonyms; transitional devices appropriate/effective; cohesive devices used to further the progression of ideas in a manner clearly relevant to the overall meaning Coherent and convincing to reader; uses transitional devices/referential ties/logical connectors to create and further a particular style	Manipulates syntax with attention to style; generally error-free sentence variety; meaning clear; non-English patterns rarely evident  Mostly error-free; frequent success in using language to stylistic advantage; idiomatic syntax; non-English patterns not evident	Meaning clear; fairly sophisticated range and variety; word usage under control; occasionally unidiomatic; attempts at original, appropriate choices; may use some language nuance  Meaning clear; sophisticated range, variety; often idiomatic; often original, appropriate choices; may have distinctions in nuance for accuracy, clarity	Uses mechanical devices to further meaning; generally error-free  Uses mechanical devices for stylistic purposes; may be error-free
10	Appropriate native-like standard written English	Appropriate native-like standard written English	Appropriate native-like standard written English	Appropriate native-like standard written English	Appropriate native-like standard written English	

Source: Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing*, 8, 265–289.

## References

- Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, 78, 465–483.
- Anton, M., & DiCamilla, F. (1998). Sociocognitive functions of L1 collaborative interaction in the L2 classroom. *Canadian Modern Language Journal*, 54, 314–342.
- Beach, R. (1989). Showing students how to assess: Demonstrating techniques for response in the writing conference. In C. M. Anson (Ed.), *Writing and response: Theory, practice, and research* (pp. 127–148). Urbana, IL: National Council of Teachers of English.
- Bell, J. H. (1991). Using peer response groups in ESL writing classes. *TESL Canada Journal*, 8, 65–71.
- Braine, G. (2003). From a teacher-centered to a student-centered approach: A study of peer feedback in Hong Kong writing classes. *Journal of Asian Pacific Communication*, 13, 269–288.
- Bruffee, K. (1973). Collaborative learning: Some practical models. *College English*, 34, 634–643.
- Bruffee, K. (1978). The Brooklyn Plan: Attaining intellectual growth through peer-group tutoring. *Liberal Education*, 64, 447–468.
- Butler, S. (1981). *The bridge to real writing: Teaching editing skills*. Presented at the University of British Columbia. (ERIC Document Reproduction Service No. ED228639).
- Carson, J. G., & Nelson, G. L. (1994). Writing groups: Cross-cultural issues. *Journal of Second Language Writing*, 3, 17–30.
- De Guerrero, M. C. M., & Villamil, O. S. (2000). Activating the ZPD: Mutual scaffolding in L2 peer revision. *The Modern Language Journal*, 84, 51–68.
- Ferris, D. R. (2003). *Response to student writing: Implications for second language students*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gebhardt, R. (1980). Teamwork and feedback: Broadening the base of collaborative writing. *College English*, 42, 69–74.
- Gieve, S. (1998). Comments on Dwight Atkinson's "A critical approach to critical thinking in TESOL." *TESOL Quarterly*, 32, 123–129.
- Graner, M. (1987). Revision workshops: An alternative to peer editing groups. *The English Journal*, 76, 40–45.
- Hansen, J., & Liu, J. (2005). Guiding principles for effective peer response. *ELT Journal*, 59, 31–38.
- Hyland, F. (2000). ESL writers and feedback: Giving more autonomy to students. *Language Testing Research*, 4, 33–54.
- Kamimura, T. (2006). Effects of peer feedback on EFL student writers at different levels of English proficiency: A Japanese context. *TESL Canada Journal*, 23, 12–39.
- Krashen, Stephen. (1982). *Principles and practice in second language learning and acquisition*. Oxford: Pergamon.
- Lee, I. (1997). Peer reviews in Hong Kong tertiary classroom. *TESL Canada Journal*, 15, 58–69.
- Lockhart, C., & Ng, P. (1995). Analyzing talk in ESL peer response groups: Stances, functions, and content. *Language Learning*, 45, 605–655.
- Lynch, B. K., & McNamara, T. F. (1998). Using G-theory and Many-facet Rasch measurement in the development of performance assessments of the ESL speaking skills of immigrants. *Language Testing*, 15, 158–180.
- Mangelsdorf, K. (1992). Peer reviews in the ESL composition classroom: What do the students think? *ELT Journal*, 46, 274–284.
- Marcus, H. (1984). The writing center: Peer tutoring in a supportive setting. *The English Journal*, 73, 66–67.
- Matsuhashi, A., Gillam, A., Conley, R., & Moss, B. (1989). A theoretical framework for studying peer tutoring as response. In C. M. Anson (Ed.), *Writing and response: Theory, practice, and research* (pp. 293–316). Urbana, IL: National Council of Teachers of English.
- McMurry, A. I. (2004). *Preparing students for peer review*. Unpublished master's project. Provo, UT: Brigham Young University.
- Mendonça, C. O., & Johnson, K. E. (1994). Peer review negotiations: Revision activities in ESL writing instruction. *TESOL Quarterly*, 28, 745–769.
- Min, H. T. (2005). Training students to become successful peer reviewers. *System*, 33, 293–308.
- Min, H. T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing*, 15, 118–141.
- Nassaji, H., & Swain, M. (2000). A Vygotskian perspective on corrective feedback in L2: The effect of random versus negotiated help on the learning of English articles. *Language Awareness*, 9, 34–51.
- Nelson, G. L., & Murphy, J. M. (1992). An L2 writing group: Talk and social dimension. *Journal of Second Language Writing*, 1, 171–193.
- O'Donnell, C. (1980, March). *Peer editing: A way to improve writing*. Paper presented at the meeting of the combined annual meeting of the Secondary School English Conference and the Conference on English Education, Omaha, NE.
- Park, T. (2004). An investigation of an ESL placement test of writing using multi-faceted Rasch measurement. *Teachers College, Columbia University working papers in TESOL and applied linguistics*, 4, pp. 1–21.
- Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing*, 8, 265–289.
- Ramathan, V., & Atkinson, D. (1999). Individualism, academic writing, and ESL writers. *Journal of Second Language Writing*, 8, 45–75.
- Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT Journal*, 59, 23–30.
- Sager, C. (1973, November). *Improving the quality of written composition through pupil use of rating scale*. Paper presented at the annual meeting of the National Council of Teachers of English, Philadelphia, PA.
- Schaffer, J. (1996). Peer response that works. *Journal of Teaching Writing*, 15, 81–90.
- Sullivan, K., & Lindgren, E. (2002). Self-assessment in autonomous computer-aided second language writing. *ELT Journal*, 56, 258–266.
- Swain, M., Brooks, L., & Tocalli-Beller, A. (2002). Peer-peer dialogue as a means of second language learning. *Annual Review of Applied Linguistics*, 22, 171–185.
- Tang, G. M., & Tithecott, J. (1999). Peer response in ESL writing. *TESL Canada Journal*, 16, 20–38.
- Teo, A. K. (2006). Social-interactive writing for English language learners. *The CATESOL Journal*, 18, 160–178.
- Thompson, C. (2002). Teaching critical thinking in EAP courses in Australia. *TESOL Journal*, 11, 15–20.

Vygotsky, L. S. (1986). *Thought and language*. London: Longman.

Warwick, P., & Maloch, B. (2003). Scaffolding speech and writing in the primary classroom: A consideration of work with literature and science pupil groups in the USA and UK. *Reading Literacy and Language*, 37, 54–63.

Winsteps (2008). *Winsteps, facets: Rasch analysis software*. Retrieved September 2, 2006, from [www.winsteps.com](http://www.winsteps.com).

**Kristi Lundstrom** recently graduated from the TESOL Master's program at Brigham Young University.

**Wendy Baker** is an assistant professor in the Department of Linguistics and English Language at Brigham Young University.