Training students to become successful peer reviewers

Hui-Tzu Min *

Department of Foreign Languages and Literature, National Cheng Kung University,
1 University Road, Tainan 70101, Taiwan

Received 5 August 2004; received in revised form 26 October 2004; accepted 16 November 2004

Abstract

Vague feedback and misinterpretation of writers’ intentions on the reviewers’ part have been found to be two major reasons why most of their comments are disregarded during writers’ revision in one EFL writing class. To resolve this problem, a training was conducted to coach these students to generate more specific comments. Four characteristics of comments that were found to facilitate students’ revisions in previous research were identified and used as guidelines during training: Clarifying writers’ intentions, identifying problems, explaining the nature of problems, and making specific suggestions. Two teacher–student conferences with each reviewer were also held to provide individual assistance. Both quantitative and qualitative analyses of students’ comments after training were conducted. Students were able to generate significantly more comments containing two or three afore-mentioned characteristics and were able to produce more relevant and specific comments on global issues. As reviewers students benefited from this training in skill improvement, confidence build-up, language acquisition and metacognitive strategy use. As writers they were able to approach topics of interest to them from multiple perspectives and to increase their vocabulary repertoire after training.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Peer review; Written feedback; EFL writers; Training studies
1. Introduction

The use of peer review in process-oriented instruction can find its theoretical support in two different but closely related disciplines: learning and rhetorical theories. In terms of learning, Vygotsky’s (1962, 1978) theory on learning and language coincides with the use of peer reviews. Vygotsky deemed social interaction an essential element for cognitive learning and accorded great importance to language in human’s thought development. To him, learning is a cognitive activity that takes place in social interaction. By the same token, writing is a learning activity in which the writer learns best through interacting with his peer reviewers. Peer review provides opportunities for literacy learning because reviewers and writers vary in their strengths, preferred modes of expression, and levels of competence. Partners with differing skills and competencies can therefore provide each other with the skilled assistance needed to extend the others’ writing competence. According to this theoretical perspective, development will occur during peer review when partners have different areas of competence and interact positively in oral or written communication that includes questioning, providing elaborated responses, and instructing.

With regard to rhetorical theories, those stressing the social nature of writing have also favored the use of peer reviews (Berlin, 1987; Bruffee, 1984, 1993; Harris, 1990). Bruffee (1984) contended that writing is derived from the “conversation” among writers in their discourse community. He argued that the collaborative environment created in peer groups is conductive to addressing high-order composition issues among writers, such as focus and idea development (cited in Stanley, 1992). Therefore, it is important to offer students opportunities to immerse themselves in constructive conversation about writing.

Despite its numerous beneficial effects (Mangelsdorf, 1992; Mittan, 1989; Tsui and Ng, 2000), criticism on peer response has also accrued. One criticism has been leveled at students’ inability to provide concrete and useful feedback. A number of researchers (Chou, 1998; Leki, 1990; Mendonca and Johnson, 1994; Lockhart and Ng, 1993; Tsui and Ng, 2000) noted that students tended to give rubber stamp advice when reviewing peers’ essays. Some attributed this to students’ lack of knowledge and skills needed for peer review and called for teachers’ intervention (Flynn, 1982; George, 1984; Wiener, 1986). To date, few researchers have answered this call by examining the efficacy of teacher assistance (Sommers and Lawrene, 1992; Stanley, 1992; Zhu, 1995) in peer response groups among native speakers of English (NSE) and learners who study English as a second language (ESL).

1.1. Literature review

Sommers and Lawrene (1992) were interested in the floor-holding between genders and examined the impact of teacher intervention on women’s talk in peer response groups. Their qualitative analysis revealed obvious gender-related differences. In teacher-directed groups, female students contributed to discussions almost as equally as their male counterparts; whereas in student-directed groups, they were not as active and were interrupted more when they talked.
Stanley (1992) examined the effect of lengthy training sessions on peer evaluation in university ESL freshman composition classes. The training sessions lasted 7 hours, during which students engaged in role-play, analyzing the genre of peers’ essays and discovering rules for effective communication. Results show that a lengthy training enhanced the effectiveness of peer evaluation. The coached group generated substantially more feedback in a more tactful way than the control group.

Zhu (1995) also investigated the effect of training on peer interaction in peer response groups in university freshman composition classes where all subjects were native speakers of English. Zhu (1995) employed a conference method to train students. The trainer met students in groups of three, discussing strategies for giving effective feedback on peer writing three times during the semester. Results show that students receiving training generated more feedback and engaged in more active interaction and negotiation than those only watching video demonstration.

1.2. Critique

A perusal of the training studies reveals that all focused on training NSE or ESL learners to function well in peer response groups rather than in paired peer review. Peer response groups differ from paired peer review not only in group constitution but also in the nature of tasks. Generally speaking, providing oral feedback is easier for NSE/ESL learners than for most EFL writers who are not used to orally expressing their ideas in a foreign language, especially under time pressure. Consequently, most instructors of EFL writing prefer using peer review in their classes. Due to the difference in task demands, issues characterizing peer response groups like floor-taking (Sommers and Lawrene, 1992) and passive interaction (Zhu, 1995) are inapplicable to EFL peer review contexts where students silently read their peers’ essays and make written comments. On the other hand, problems characterizing paired peer review in EFL contexts like misunderstanding the writer’s intentions and offering vague comments (Min, 2003) seem to be a peripheral problem for most NSE/ESL students. Previous NSE/ESL training studies presumed students’ capability to understand writers’ intentions and provide specific feedback. None underscored the importance of reviewers’ clarifying writers’ intentions, explaining to writers why they think there is a problem in the text, and providing specific feedback based on the said explanation. Stanley (1992) noted a tendency among ESL students to “overinterpret” their partners’ essays. The problem that lies with EFL peer reviewers, is, however, more than overinterpretation. It is an aggregate of over- and mis-interpretation as well as a lack of skills in providing specific feedback.

2. Rationale for the current study

Given the difference in tasks, strategy training centering on searching for tactics to convey opinions in NSE/ESL studies appears to lack immediate relevance to EFL students who are struggling to understand writers’ intentions and provide specific
written feedback. A different kind of training designed to address the issues of EFL paired peer review is needed. This kind of training must be based on a thorough understanding of the problems EFL students encounter during peer review.

An earlier analysis of why most peer comments made by EFL university sophomore students failed to be incorporated shows that misunderstanding of writers’ intentions and vague comments were two major causes (Min, 2003). Drawing on these findings, the researcher designed a four-step procedure for paired peer review: Clarifying writers’ intentions, identifying problems, explaining the nature of problems, and making suggestions by giving specific examples. These steps were drawn on an inductive analysis of diverse oral communicative behaviors found to facilitate in shaping peer revision in previous training studies (Stanley, 1992) and of those examining peer negotiation (Mendonca and Johnson, 1994; Tang and Tithecott, 1999; Villamil and De Guerrero, 1996) as well as the researcher’s teaching experience (see Appendix A).

A step-by-step procedure may appear rigid and unnecessary in peer response groups where ideas and doubts can be exchanged and clarified during face-to-face interaction. However, it is crucial for EFL paired peer review since there is usually little time for the reviewer and writer to discuss the written comments in class due to time constraints. Furthermore, the use of paired peer review is preferred by most EFL students because the written mode allows them more time to organize their ideas in English. It also helps the writers because they could review the comments as many times as they want while making revisions.

3. Research questions

Four research questions are addressed in this study: After training, (1) do students provide a greater amount of relevant and specific feedback? (2) does the number of comments that followed more steps increase? (3) do students provide a greater amount of feedback on the global features of peer writing? (4) how do students benefit? Most ESL training studies on peer review urge students to focus on idea development and organization of content and de-emphasize the importance of attending to word usage, grammar and spelling. Nonetheless, Ferris (1997) found that students also welcome comments that require grammar accuracy and are more likely to revise their drafts upon receiving such comments. Given that the ultimate aim of using peer review is to help improve students’ compositions, the researcher placed an equal premium on both global and local issues while coaching students.

4. Methodology

4.1. Subjects and the peer review

Participants were 18 EFL sophomore students in the researcher’s composition class at a large university in southern Taiwan. All were English majors with interme-
diately English proficiency, who were introduced to the practice of peer review at the beginning of the fall semester.

In this writing class, students were required to search information and quote it to substantiate their opinions in their expository essays. Despite the researcher’s provision of a guidance sheet (see Appendix B) during peer review for the first essay, most peer comments were perfunctory, made only to answer the teacher’s questions on the guidance sheet. Some students even answered questions with a Yes or No. Seeing the students’ misinterpreting the guidance sheet as a series of questions for them to answer and their confusion about how to make comments, the researcher embarked on a training session.

4.2. The training

The training session was composed of two phases. The first phase was an in-class demonstration and modeling which lasted 2 hours per week for a total of 4 hours. During the in-class training, she taught students how to make comments on essays composed by former students by modeling to them the four-step procedure. Techniques for implementing each step were also modeled in class. After the demonstration, the students were required to do peer review for the second and third essays in class.

The second phase was two 30-minute teacher–student conferences outside of class. After the 2nd and 3rd peer review, the researcher collected reviewers’ comments and checked them carefully. Then she held two 30-minute conferences (one after each peer review) with each reviewer. The conference approach was selected primarily due to its pedagogical benefits (Harris, 1990). Teacher–student conferences not only help teachers establish rapport with their students and tailor their instruction to meet students’ needs but also allow teachers to check students’ comprehension of their instruction and feedback. In addition, the conference approach can provide both “substantive” and “procedural” facilitation (Scardamalia and Bereiter, 1985). During the conferences, the researcher offered “substantive” assistance by working with students on modifying comments that fail to follow the prescribed steps of the procedure and by responding to their revision of these comments. She also provided “procedural” support by reminding students of the questions on the guidance sheets and alerting them to the next step of the procedure through use of various oral prompts (e.g., “Does the third paragraph begin with a topic sentence?”).

4.3. Procedure

After the four-hour in-class training and 18 hours of teacher–student conferences outside class, the researcher required the students to compose the last essay (The advantages/disadvantages of —— (a new technical invention)) at home and bring their drafts to class for peer review. The students were allowed two full hours to provide written comments on two of their classmates’ compositions in class with the aid of the guidance sheet. After the peer review, the researcher collected their comments and carefully compared them with the peer review feedback
generated prior to training, that is, comments produced during the peer review for the first expository essay “An Analysis of a Cross-Cultural Misunderstanding”. Although these two topics appear to be tapping students’ different knowledge bases, they are comparable in that both required the students to use their analytical ability to either explain the sources of cultural misunderstandings or the advantages/disadvantages of a certain technological invention. The average number of words produced for the first and last compositions were 263 and 257 words respectively.

4.4. Analysis

Both quantitative and qualitative methods were used. The quantitative analysis was mainly a text analysis, comparing the number of comments, words, the number of comments on global and local issues, and the number of steps each comment contained before and after training. Paired t tests were run to see if there was any significant mean difference in the afore-mentioned items. Regarding the qualitative analysis, the researcher analyzed students’ reactions to this training in their journal entries as well as the tone and language in reviewers’ comments. The students were required to turn in five journal entries for the writing class. The original purpose of journal writing was to provide students with another means of private dialogue with the researcher regarding their opinions on the writing class. The students were encouraged to discuss anything related to the writing class, including their opinions on peer review training.

4.5. Coding

For the quantitative analysis, two trained independent raters first tallied the number of written comments and words for the first and fourth essays to obtain a general picture of the amount of feedback generated before and after training. Next, they classified the comments into two categories—one related to global issues (idea development and organization) and the other local problems (word usage, grammar, spelling, and punctuation). Then, they coded each comment in brackets according to the prescribed four criteria: clarifying writers’ intentions [clarify], identifying problems [problem], explaining the nature of the problems [explain], and making specific suggestions [suggestion]. Interrater reliability for the number of words, comments, categories, and four steps are 100%, 100%, 97%, and 93%, respectively.

Examples of applying this coding scheme to peers’ comments are presented below.

Comment 1: “call for” means to need a particular action or behavior [explain]. Did you mean that? [clarify] If not, try to think a more appropriate word to replace it. [suggestion]”

Comment 2: I think on this point, the description of the two cultures is not parallel [problem]. (Chinese = the definition of chingkeh (treat friends); America = the way they split bills [explain]).
Since the first reviewer discussed the usage of words, her comment was classified under the category of local issues. The second reviewer focused on idea development. So her comment was categorized under global issues. For the qualitative analysis, the two raters read the students’ responses in their journals and categorized them into negative and positive feedback and how they benefited from this training.

5. Findings

5.1. Number of comments

Table 1 shows a general picture of the size of mean differences in comments produced before and after the training. As clearly demonstrated, the numbers of comments ($t = -2.741, p < .05$) and words ($t = -3.418, p < .05$) produced post-training were significantly higher than those prior to training. Although a greater amount of feedback does not guarantee more relevant and specific feedback, it suggests that the reviewers were more engaged in peer review than they had been before.

5.2. Number of comments following the steps in the procedure

To examine whether the comments produced post-training were more specific than those generated before, the researcher further analyzed each comment to examine how many steps of the procedure each comment contained. Remember that the students were strongly encouraged to follow the four steps during peer review. If students included more steps in their comments, such feedback was considered more relevant and specific because each step was properly sequenced to address one aspect of the issue.

The results show no significant mean difference in the number of comments that contained only one step ($t = -1.026, p > .05$) before and after training. However, there were significant mean differences in the number of comments that included two ($t = -5.276, p < .05$) and three ($t = -6.028, p < .05$) steps. These findings suggest that while students still generated the same amount of comments that contained only one step, the number of comments containing more steps increased significantly. This phenomenon is an effect of the training. In fact, Table 2 shows that prior to

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Mean differences in number of comments and words before and after the training</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Comments</td>
<td>$-3.8611$</td>
</tr>
<tr>
<td>Words</td>
<td>$-76.450$</td>
</tr>
<tr>
<td>One step</td>
<td>$-37.9278$</td>
</tr>
<tr>
<td>Two steps</td>
<td>$-2.0833$</td>
</tr>
<tr>
<td>Three steps</td>
<td>$-0.8500$</td>
</tr>
</tbody>
</table>

* $p < .05$. 
training, the comments that contained only one step outnumbered the total of two and three steps \((t = 2.929, p < .05)\). However, after training, the difference began to shrink, due to a significant increase in the number of comments that employed two and three steps in the procedure \((t = .679, p > .05)\).

### 5.3. Number of comments on global issues

#### 5.3.1. Comparison between global and local comments after training

Of the 138.37 comments made on the first essay before training, 53% were related to global features, and 47% to local issues. After training, 57% of the 192.9 comments made on the last essay belonged to global issues, and 43% to local features. As demonstrated in Table 3, the means of global comments, both before and after the training, were larger than those of local comments. The finding indicates that students were able to comment more on global issues both prior to and after training. Although the mean differences between global and local comments did not reach statistical significance post-training, such differences became larger \((M = .4461\) before

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. deviation</th>
<th>Std. error mean</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step1 vs. Steps 2 and 3 (before the training)</td>
<td>2.6556</td>
<td>3.8463</td>
<td>0.9066</td>
<td>2.929</td>
<td>0.009*</td>
</tr>
<tr>
<td>Step1 vs. Steps 2 and 3 (after the training)</td>
<td>0.6333</td>
<td>3.9583</td>
<td>0.9330</td>
<td>0.679</td>
<td>0.506</td>
</tr>
</tbody>
</table>

* \(p < .05\).

### Table 3

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. deviation</th>
<th>Std. error mean</th>
<th>t</th>
<th>Sig. (2 tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global1(^a)</td>
<td>4.0667</td>
<td>2.3029</td>
<td>0.5428</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local1(^b)</td>
<td>3.6206</td>
<td>3.0702</td>
<td>0.7237</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global2(^c)</td>
<td>6.000</td>
<td>2.1693</td>
<td>0.5113</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local2(^d)</td>
<td>4.7167</td>
<td>4.2922</td>
<td>1.0117</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global1/Local1(^e)</td>
<td>0.4461</td>
<td>3.2310</td>
<td>0.7615</td>
<td>0.586</td>
<td>0.566</td>
</tr>
<tr>
<td>Global2/Local2(^e)</td>
<td>1.2833</td>
<td>5.0885</td>
<td>1.1994</td>
<td>1.070</td>
<td>0.300</td>
</tr>
<tr>
<td>Global1/Global2(^e)</td>
<td>−1.9333</td>
<td>2.9462</td>
<td>0.6944</td>
<td>−2.784</td>
<td>0.013*</td>
</tr>
<tr>
<td>Local1/Local2(^e)</td>
<td>−1.0961</td>
<td>5.0838</td>
<td>1.1983</td>
<td>−.915</td>
<td>0.373</td>
</tr>
</tbody>
</table>

\(^a\) Global1 means the number of comments on global issues before the training.

\(^b\) Local1 means the number of comments on local issues before the training.

\(^c\) Global2 means the number of comments on global issues after the training.

\(^d\) Local2 means the number of comments on local issues before the training.

\(^e\) \(p < .05\).
training, $M = 1.2833$ after training), suggesting that students were able to provide a greater amount of feedback on the global features after training. This might be an effect of training.

5.3.2. Comparison of global comments before and after training

Furthermore, the mean difference in the number of comments on global issues before and after training reached significance ($t = -2.784$, $p < .05$). This indicates that students tended to allocate more attention to macro issues such as idea development and organization post-training.

5.4. Students’ reaction to training

5.4.1. As a reviewer

All 18 students expressed positive comments in their journal entries about this training. They all pointed out that the four-step procedure helped them become better reviewers, although following the four steps was both time- and energy-consuming. Students benefited from this training in four different aspects—skill improvement, language acquisition and self-monitoring and confidence building. The first positive effect of this training is that all students learned to provide more feedback to global issues. As succinctly pointed out in one student’s entry: “I realized that the most important thing of composing is ideas and organization, not vocabulary or grammar. I learned how to give suggestions to others by following the steps in the procedure and showed my own opinions, and saw if others also agreed with my thoughts.”

The second benefit most students reaped is language acquisition. Many students (13 out of 18) expressed that making specific suggestions helped them increase their vocabulary repertoire: “When I tried to explain why I thought a certain part was problematic, I had to look words up in the dictionary sometimes. At first, it was quite a nuisance. But later I found that my vocabulary increased a lot. I think it’s an advantage of following the steps.”

The third benefit is self-monitoring. Eleven students mentioned in their journals that this training made them reflect on their own problems and seek out solutions for themselves. “This training indeed helped me revise my writing. When I read others’ essays, I discovered the same mistakes I made in my own composition. It would have been difficult, if not impossible, for me to find my own mistakes if I hadn’t read the same ones in others’ essays.”

With regard to the affective impact, the training helped less-advanced reviewers gain confidence in viewing themselves as competent readers. It helped them realize that incomprehension of the writer’s intentions was not always an indication of their limited English proficiency. Ten students mentioned this affective boost in their journals. A change of mentality is succinctly summarized in one student’s entry, “during the first peer review, I did not know what to do when I did not understand the writer. I thought it was my poor English ability that made me unable to understand him. And if I told him that I didn’t know what he was saying, I would feel embarrassed. So I pretended that I understood him, although I had lots of questions on my mind. But during the teacher–student conference, I learned that it was not necessarily my
fault when I did not understand the writer’s intention. The writer should write in a way that’s comprehensible to me, the reader. All I needed to do was to ask him what he meant and discussed with him my interpretation of his intention.”

5.4.2. As a writer

Fifteen students thought that the training helped them as writers. They learned from their peers how to focus their ideas and view things from different perspectives. Such help assisted them in elaborating their ideas in a much more specific way. As succinctly commented by one writer, “before training, I usually got vague suggestions from my classmates and didn’t understand why they considered a certain part of my writing problematic. After training, my classmates were able to explain my problems in a much clearer way and give me more specific suggestions. Their suggestions really helped me refine my ideas a lot. When I looked back at my first draft (of the last essay) and compared it with the second draft, I found that the later revision was much more refined and improved, thanks to my classmates’ helpful feedback.”

In addition to helping writers refine and elaborate their ideas, 15 students also thought that their classmates’ comments provided different perspectives on the same issue, helping broaden their horizons and enrich their content. “Some classmates give me good suggestions, which give me a new way to think. We always have our own blind spot. The outsider [reviewer] is always clear.”

5.5. Change of language and tone

Besides self-report benefits, the researcher also found two other qualitative changes in students’ comments after training—language and tone. The language used prior to training was short and impersonal, rendering the tone like teachers’ corrections over students’ assignments (indicated by italics). Examples abounded in comments on macro and local issues. An illustration of this commanding tone can be found in a comment on a writer’s organization, “You should put some phrases before you make this quotation because the last paragraph is unrelated to the fourth paragraph.”

After training, however, the comments sounded like readers communicating ideas with writers and thus the language longer, the tone much friendlier. On the surface, the difference in language and tone may be just a replacement of modals. However, such replacements also reflect a subtle change of students’ mentality and their appropriate use of language to convey this attitude. Prior to training, some students considered peer review to be a chance to play the role of teacher. After the researcher’s explanation during teacher–student conferences, these students began to take on a new stance. They started to communicate their ideas from a reader’s perspective rather than from a teacher’s role. They empathized with the writer and provided feedback from the writer’s viewpoint. For example, a reviewer made the following comment on idea development after training, “If you’re trying to say many people have more than one cell phone, maybe you can say it in this way...” It is through this attitudinal change—from a criticizing attitude to an empathizing one—that the comments are made more helpful and valuable to the writers.
6. Discussion

The current study investigated the effect of coaching students to become better peer reviewers. A four-step procedure was taught to students when they performed peer review. During the training session, two teacher–student conferences were also held to address individual problems when providing feedback. The quantitative analysis revealed that the students were able to produce significantly more comments post-training. The number of words for single, two, and three steps respectively was significantly higher than that produced before training. These findings demonstrate that reviewers’ feedback was much more relevant and specific because it revolved around clarifying, identifying and expounding a single issue and suggesting a way to improve it. In addition, the students also made more comments on global issues both before and after training.

The qualitative analysis revealed reviewers’ and writers’ positive attitudes toward this training. All reviewers appreciated this training because they benefited from it methodologically, linguistically, cognitively, and psychologically. It helped them focus on commenting global issues, increase their vocabulary, consciously avoid making the same errors spotted in their classmates’ compositions, and boost their self-confidence as competent readers. Not surprisingly, all writers reaped the fruit of the training too because they learned how to refine their descriptions of certain issues in different expressions and to approach an issue from different perspectives.

6.1. Increasing number of comments: Effect of training or reaction to writings of different qualities

The study found an overall increased quantity of comments and words after the training. But one question remains. How valid and meaningful is it to serve as a measure of improvement in giving feedback, as especially the post-training feedback comments are related to reviews of new essays, not revised drafts of earlier essays. Some might argue that the increased quantity of feedback was simply a response to the quality of essays than to the training. However, this argument failed to take into account the students’ profile. As is known, a prerequisite of responding to an essay is that reviewers know what to attend to and how to respond to it. Most of the students in this study did not possess such knowledge or skills prior to training. Comments on students’ perceived difficulty in providing feedback abounded in their journals before training: “I didn’t know how to make comments after reading the essay many times.” “During the first peer review, I had no idea about how to give feedback to my classmates. Therefore, I started to revise his or her grammar mistakes.” It was not until the beginning of the training that such perceived difficulty started to lessen. “After the training and conference with Amy (the teacher), I began to have a more concrete idea of what I was supposed to do during peer review.” “I learned to pay attention to ideas and organization, not just grammar.” “At first, I thought that my suggestions were minor and unimportant so I didn’t point them out. After talking to Amy, I found I was wrong. Now I think every suggestion has its own value. Every point is deserved to be discussed as long as I have a question.
about it.” Given these comments, it appears that the underlying reason for fewer peer comments before training was due to students' inability to provide feedback, not an effect of responses to different writing qualities. The last two excerpts directly point to the effect of the training, which contributed to a significant increase of comments after training.

The training effect was also manifested in the number of steps each comment included. Prior to training, most reviewers gave rubber stamp advice which included only one step. After training, however, the number of comments containing two or three steps significantly increased. Students' journal entries also corroborated such a finding.

6.2. Necessity of clarifying writers' intentions

None of the comments after training contained all four steps. A closer examination of the types of comments revealed that not all comments started with clarifying the writers' intentions. In fact, many students expressed that, having frequently worked with the same partners throughout the semester and shared the same language and cultural background, they usually understood what their partners wanted to convey. Thus they did not bother with identifying writers’ intentions when they located a problem. This raises a question. Should this step be included in the training procedure? Based on the previous findings (Min, 2003; Tang and Tithecott, 1999) that peers were likely to misunderstand writers’ intentions, leading to an irritated feeling on the writers’ part (Tang and Tithecott, 1999), the researcher still deems this step a requirement, especially in ESL writing classes where students are from diverse cultural and language backgrounds and in EFL classrooms where instructors always encourage students to change partners. This step is mainly to remind readers that they need to be certain about writers’ intentions before proceeding to the following three steps in the procedure. It also assures less advanced readers of their right to ask writers to express their intentions in a way that is comprehensible to them. It can alleviate the uneasy feelings among readers with limited English proficiency, who might attribute their inability to comprehend writers’ intentions to their low language proficiency, and thus dare not raise questions for fear of losing face (Carson and Nelson, 1996). Furthermore, this step is also instrumental to developing among EFL writers an awareness of their responsibility in the written communication of the Western world (Hinds, 1987). They need to know that in Western written communication writers should be more responsible for explicitly expressing their intentions. If there is a communication gap, writers should try harder to bridge the gap rather than leave the readers to gauge their intentions. Finally, clarifying writers’ intentions also helps writers become more aware of the audience's needs. Consequently, they would generate more reader-friendly prose (Flower and Hayes, 1981).

6.3. Comments on global/local features

Besides employing more steps in providing feedback, the students also made more comments on global features than on local ones both before and after training. This
shows that these EFL students were able to attend to idea development and paragraph organization when doing peer review. In the meantime, they also attended to local issues because the number of comments on vocabulary and grammar also increased after training. This result can be deemed as an effect of the teacher’s equal emphasis on global and local issues, given that she did not discourage the students from making comments on micro issues.

Many ESL writing instructors and researchers contend that students should pay more attention to macro issues during peer review. They ask students to focus on idea development and organization rather than on grammar and spelling and consider peer review unsuccessful if students fail to do so (Leki, 1990). Such an emphasis, of course, is to wean students off their old habit of attending to only micro-level issues. However, deliberate overlooking of grammatical errors and word usage might send an unwanted message that correct usage of grammar and words is not important as long as their ideas are fine. This attitude toward minor language problems might be suitable with ESL writers who generally do not have serious problems expressing their ideas in English. But it does not rest well with most EFL writers who are constantly grappling with English in expressing their ideas while composing academic essays. These students need input on both content and form so that they can generate writings that are rich and organized in content and acceptable in form.

The researcher believes that EFL writing instruction, including peer review, should assist students in developing and organizing ideas, and expressing them in appropriate language.

6.4. Sociocognitive benefits of peer review

Sociocultural theory emphasizes that development requires interaction and the presence of support from a more skilled adult or partner. This finding suggests that with proper training, student reviewers exhibited many constructive patterns of interaction and were capable of providing scaffolding (assistance) that is considered critical for development to writers. Sociocognitive activities such as reading, questioning, pointing to trouble sources, discussing ideas and scaffolding activities including explaining, instructing, giving specific comments abounded in peers’ feedback. Writers also reported reaping benefits from peer review such as gaining multiple perspectives on a single issue and language acquisition. They relied on reviewers as resources when they had questions about content, organization and language. All expressed that they could not have reaped such benefits if they had not corroborated with their peers, a testimony to Vygotsky’s theory of sociocognitive learning.

7. Conclusion

Some teachers and researchers have reservations about practicing peer review in EFL contexts, although they agree with its underlying philosophies. Their major
concerns revolve around technical issues, which can be resolved by proper intervention. This study provides reassuring evidence that students do learn to become better peer reviewers if they are provided with proper training and individual assistance over a period of time. Although the results of this study cannot be generalized to other contexts due to a lack of control group and the small number of students involved, interested writing instructors and researchers can test the strength of the four-step training procedure in experimental studies. More research also needs to compare the number of comments writers incorporate into their revisions before and after training to validate writers’ self-reports and to examine the impact of peer review on revision quality.

Appendix A. Definitions and examples of the four training steps

<table>
<thead>
<tr>
<th>Step</th>
<th>Definition</th>
<th>Behavior with similar intentions reported in other studies</th>
<th>Examples of comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Clarifying the writer’s intention</td>
<td>Reviewers try to get further explanation of what writers have said or what is not clear to them in the essays (e.g., an unknown term, an idea)</td>
<td>Pointing (Stanley, 1992) Request for explanation (Mendonca and Johnson, 1994) Requesting clarification (Tang and Tithecott, 1999)</td>
<td>“What do you mean by ‘college-graduate society freshmen?’”</td>
</tr>
<tr>
<td>2. Identifying the problem</td>
<td>Reviewers announce a problematic word, phrase, sentence or cohesive gap</td>
<td>Trouble sources (Tang and Tithecott, 1999; Villamil and De Guerrero, 1996)</td>
<td>“I think on this point, the description of the two cultures is not parallel.”</td>
</tr>
<tr>
<td>3. Explaining the nature of the problem</td>
<td>Reviewers explain why they think a given term, idea, or organization is unclear or problematic, which should or should not be used in the essay</td>
<td>Explanations (Mendonca and Johnson, 1994)</td>
<td>“You should put some phrases before you make this quotation because the last paragraph is unrelated to the fourth paragraph.”</td>
</tr>
</tbody>
</table>
Appendix A (continued)

<table>
<thead>
<tr>
<th>Step</th>
<th>Definition</th>
<th>Behavior with similar intentions reported in other studies</th>
<th>Examples of comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Making specific suggestions</td>
<td>Reviewers suggest ways to change the words, content, and organization of essays</td>
<td>Suggestions, Grammar corrections (Mendonca and Johnson, 1994) Advising, Collaborating (Stanley, 1992) Giving directives, Instructing (Villamil and De Guerrero, 1996; Tang and Tithecott, 1999)</td>
<td>“If you’re trying to say many people have more than one cell phone, maybe you can say it in this way → 'The majority of people have a cell phone with them, some even with more than one.’”</td>
</tr>
</tbody>
</table>

* This example can be coded as an attempt both at explaining the nature of the problem and making suggestions.

Appendix B. Guidance sheet

1. Read the first sentence. What is the topic? What is the controlling idea? Circle them. Is the topic sentence a statement of opinion, intent, a combination of both, or just simple fact? If it is a statement of fact, help the writer rewrite it so that it becomes a real topic sentence (i.e., a statement of opinion, intent, or a combination of both).
2. After reading the topic sentence, what do you expect to read in the following sentences?
3. Now read the following two or three sentences. Did the writer write according to your expectation(s)? If not, what did the writer write instead? Do you think that writer was sidetracked? Go back to the bridge (second sentence). Did the author choose a word that is not the controlling idea to develop? Did the author talk about an idea more general than or in contrast to the controlling idea? If none of these applies, reread the topic sentence to make sure that you understand the writer’s intention.
4. Read the examples. How many examples are there? Are they well balanced (in terms of sentence length and depth of discussion)? Are they relevant to the controlling idea in the topic sentence? If not, explain to the writer why they are irrelevant. Also work with the writer to think of more things to talk about if the examples are too general or to delete some of the redundant sentences.
5. Read the last few sentences in the paragraph. Is there a restatement at the end of the paragraph? If not, work with the writer on a concluding sentence.
6. What did you learn from reading this paragraph, either in language use or content? Is there anything nice you want to say about this paragraph? Are there any grammatical errors or inappropriate word usage?
References


