The effects of trained peer review on EFL students’ revision types and writing quality

Hui-Tzu Min*

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

Abstract

This preliminary classroom study aims to examine the impact of trained responders’ feedback on EFL college students’ revisions, both in terms of revision types and quality. After a 4-hour in-class demonstration and a 1-hour after-class reviewer-teacher conference with each student (n = 18), the instructor/researcher collected students’ first drafts and revisions, as well as reviewers’ written feedback, and compared them with those produced prior to training. Results show that students incorporated a significantly higher number of reviewers’ comments into revisions post peer review training. The number of peer-triggered revisions comprised 90% of the total revisions, and the number of revisions with enhanced quality was significantly higher than that before peer review training. The researcher concludes that with extensive training inside and outside of class, trained peer review feedback can positively impact EFL students’ revision types and quality of texts directly.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Peer review training; Revision quality; Revision types; EFL writing; Taiwan

1. Introduction

The practice of peer response/review has been burgeoning in both L1 and ESL/EFL writing classes since last decade, given its strong support from social learning (Vygotsky, 1962, 1978) and rhetorical theories (Berlin, 1987; Bruffee, 1993; Harris, 1990). Proponents of peer response/review have made a plethora of claims about its cognitive, affective, social, and linguistic benefits, most of which have been substantiated by extant empirical evidence. Peer response/review has been found to help both college (de Guerrero & Villamil, 1994; Mendonca & Johnson, 1994; Villamil & de Guerrero, 1996) and secondary (Peterson, 2003; Tsui & Ng, 2000) students obtain more insight into
their writing and revision processes, foster a sense of ownership of the text (Tsui & Ng, 2000), generate more positive attitudes toward writing (Min, 2005), enhance audience awareness (Mendonca & Johnson, 1994; Mittan, 1989; Tsui & Ng, 2000), and facilitate their second language acquisition (Byrd, 1994; Lockhart & Ng, 1995) and oral fluency development (Mangelsdorf, 1989).

In contrast to the large number of studies centering on the cognitive, affective, social, and linguistic benefits of peer response/review groups, few studies have examined the extent to which peer feedback is incorporated into students’ subsequent revisions (Chou, 1999; Connor & Asenavage, 1994; Mendonca & Johnson, 1994; Nelson & Murphy, 1993; Lockhart & Ng, 1993; Tsui & Ng, 2000). Results from these studies reveal a low ratio: ranging from 5% (Connor & Asenavage, 1994), 22% (Chou, 1999), less than 50% (Paulus, 1999; Tsui & Ng, 2000), to a little above 50% (Mendonca & Johnson, 1994; Tang & Tithecott, 1999). If peer feedback is instrumental to students’ writing as reported in the literature, why do the majority of peer comments fail to be utilized in students’ subsequent revisions?

One of the answers to this question, as pointed out by a number of researchers, lies in students’ inability to provide concrete and useful feedback (Chou, 1999; Leki, 1990; Mangelsdorf & Schlumberger, 1992; Mendonca & Johnson, 1994; Lockhart & Ng, 1993; Tsui & Ng, 2000). Students, both native speakers of English (henceforth NSE) and learners who study English as a second/foreign language (henceforth ESL/EFL learners), tend to give rubber stamp advice when reviewing peers’ essays. Some attributed this phenomenon to students’ lack of knowledge of and skills for peer review and called for teacher intervention (Flynn, 1982; George, 1984; Wiener, 1986). Researchers who have answered this call by examining the efficacy of teacher assistance in peer response groups among NSE/ESL learners (Stanley, 1992; Zhu, 1995) and in paired peer review among EFL learners (Min, 2005) all reported positive effects.

2. Literature review

Stanley (1992) provided lengthy training in peer evaluation to students in an ESL freshman composition class. Her training focused on familiarizing students with the genre of their classmates’ writing and introducing techniques of effective communication. Employing a conversational analysis approach to categorizing the evaluators’ responses, Stanley found that the coached groups made substantially more responses and more types of responses than the uncoached groups. A subsequent analysis of the drafts also revealed more revisions in response to peer evaluation in the coached groups than in the uncoached groups.

Zhu (1995) employed a small group conference approach to training L1 peer responders in university freshman composition classes. Both the experimental and control groups watched a demonstration video to learn some fundamental concepts about peer response. The experimental group, in addition, met with the instructors in groups of three, three times during the semester. Each teacher–student conference consisted of two phases—a read aloud by a volunteer student of his/her essay with peers reading along, followed by a discussion of the essay and suggestions for revision. During the discussion session, the instructors not only encouraged responders to critically mull over the merits and shortcomings of the essay and to provide specific suggestions but also demonstrated tactics writers could employ to illicit feedback and seek clarifications from their responders. Zhu (1995) reported that such peer response training had a significant effect on both the quantity and quality of feedback.

Min (2005) conducted a classroom study to train 18 responders in a sophomore EFL writing class. She identified four characteristics of comments reported to facilitate students’ revisions in previous research—clarifying writers’ intentions, identifying problems, explaining the nature of
problems, and making specific suggestions, and used them as guidelines during in-class training. She also employed Zhu’s (1995) conference method to meet with each responder twice to provide individual assistance. Subsequent text analyses of the written comments generated by responders post peer review training revealed that responders could produce significantly more comments containing two or three aforementioned characteristics and were able to produce more relevant and specific comments on global issues.

Taken together, NSE and ESL/EFL students coached to be successful peer reviewers were found to generate substantially more feedback in a more tactful and active negotiation than did the control group (Stanley, 1992; Zhu, 1995), and to focus more on commenting on global features in more detail (Min, 2005; Zhu, 1995). Yet, little knowledge is shed on how a higher level of engagement in and commitment to peer interaction on the reviewers’ part post peer response training influences students’ revision quality (Stanley, 1992; Zhu, 1995). Stanley (1992) touched upon this issue merely in a one-line summary: “The drafts of the coached groups evidenced more responses to peer evaluation than did the uncoached groups’ drafts (p. 229).” One did not know, however, if the quality of those drafts responding to more peer comments in the coached group was better than those responding to fewer peer comments in the uncoached one. To date, very few researchers have conducted a follow-up study to investigate the impact of responders’ training on shaping students’ revisions and the quality of their revisions (Berg, 1999; Paulus, 1999).

In an attempt to address the previous issue, Berg (1999) conducted a quasi-experimental study to investigate how trained peer response shapes ESL college/graduate students’ revisions and revision quality. The experimental group received various kinds of peer response training activities ranging from 5 to 45 min each. The control group, on the other hand, received no instruction in how to conduct peer response. A comparison of the first and revised drafts written by both groups revealed that the trained response group made significantly more meaning changes than the untrained group, and the quality of revisions made by the trained response group was significantly better than that of the untrained group, regardless of students’ L2 language proficiency. Drawing on these two findings, Berg (1999) concluded that trained peer response did exert positive impact on ESL students’ revision types and quality.

Although not exactly a follow-up study investigating the impact of peer review training on revision types and quality, Paulus’s (1999) classroom research examined the effect of peer and teacher review on 11 international students’ revisions in a pre-freshman composition class at a university. Using think-aloud protocols to identify sources of revision, Paulus found that these students made more surface changes than global ones. More than one-third of the revisions made in the second draft (32.3%) and only 1% of the revisions made in the third draft were peer influenced. Of all the 843 changes made to both drafts of the essays, peer feedback influenced 13.9% of these revisions, whereas teacher feedback influenced 34.3%. Although a subsequent analysis revealed improved quality from the first to the third draft, peer feedback, constituting approximately only 14% of the total revisions, contributed less than teacher feedback and other-influenced feedback to the enhanced quality.

3. Critique

Despite the researchers’ effort to specifically address the impact of trained peer response on revision types and quality between the measurable processes and outcomes (Berg, 1999; Paulus, 1999), two thorny questions still remain unresolved. First, are the quality-enhancing revisions made by trained ESL (Berg, 1999) response groups a direct result of the training?
Second, can trained peer review feedback alone significantly enhance revision quality? Berg (1999) failed to identify the sources of revision (peer- versus self-feedback) while reporting enhanced revision quality. Paulus (1999) traced the source of revision in students’ drafts through think-aloud protocols, yet her analysis showed that peer feedback contributed considerably less (13.9%) than teacher (34.3%) and self/other feedback (51.8%) to the enhanced revision.

ESL/EFL researchers have yet to discover “the exact characteristics of trained and untrained peer talk and their relationships to revision strategies and writing outcomes” (Berg, 1999, p. 233). To better understand the exact relationship between trained peer review and revision types and quality, it is imperative to first identify the source of revision in students’ essays and then investigate the ratio of peer-influenced revisions to the total revisions and revision quality. Until these questions are addressed, no definitive conclusion about the positive shaping impact of trained peer response on revisions can be reached.

4. Rationale for current study

A previous classroom-based study on peer review training (Min, 2005) revealed that a group of EFL students, after receiving extensive of coaching to be effective peer reviewers, was able to generate more specific and relevant written feedback on global features of their peers’ compositions. Presumably, peer feedback generated by students who have received such extensive peer review training should be more refined, relevant, and to the point than that produced by the same group of students prior to training. Drawing on this line of reasoning, the researcher, also the instructor of the writing class where the peer review training took place in the present study, aimed to investigate whether the same group of students would incorporate the trained peer review feedback into their revisions and whether the ratio of such incorporation was higher than that before peer review training. In addition, the researcher also sought to examine the issue of revision quality. She wanted to find out the ratio of peer-influenced revisions and the impact of such revisions on revision quality. The research questions addressed in the study thus included: (1) What are the rates of students incorporating peer’s written feedback into revisions prior to and post peer review training? (2) What are the ratios of students’ revisions in response to peer feedback prior to and post peer review training? (3) Do independent raters and the researcher find the revision quality improved post peer review training? (4) What kinds of feedback did student writers incorporate into their texts post peer review training? (5) What type of revisions lead to independent raters’ and the researcher/instructor’s definition of better text (see later definition)? The first two questions are related to the strength of the impact of peer feedback on students’ revisions, and the latter three seek to examine the direction (positive/negative) of this impact on revision quality.

5. Methodology

5.1. Participants

The study reported in this paper was part of a larger study (Min, 2005) investigating the effects of peer review training in one EFL writing class in an urban university in southern Taiwan. Participants were 18 sophomores in the instructor/researcher’s writing class. There were 16 females and 2 males, and their average age was 19. All were native speakers of Mandarin Chinese and had passed the Intermediate Level English Test of the General English Proficiency Test.
(GEPT) administered by the Language Training and Testing Center in Taiwan before being admitted to the Department of Foreign Languages and Literature at the university. Their English proficiency was approximately between 523 and 550 on the Test of English as a Foreign Language (TOEFL) exam. None of them had peer review training prior to the study.

5.2. Setting

The students in this classroom study were English majors enrolled in a year-long writing course, the focus of which was to develop their expository essay writing skills. This course met two times a week for 18 weeks each semester.

The instructor adopted a modified “writing cycle” (Tsui & Ng, 2000) in designing her writing class. The whole cycle was sequenced as follows (see Fig. 1): brainstorming, writing the first draft, written peer feedback, writing the second draft, oral presentation and peer oral response, teacher–writer conference (both oral and written comments) on the second draft, writing the third draft, teacher’s written feedback to the third draft, and writing the fourth draft (final) draft.

Students were required to search for information and quote it to substantiate their opinions in their essays in the second semester. Four Topics were written during the course of the semester—“The advantages/disadvantages of ... (a new technical invention)”, “How to prevent

Fig. 1. The writing cycle.
environmental pollution”, “Factors contributing to X’s success”, and “An analysis of an incident of cross-cultural misunderstanding”. Despite the instructor’s provision of guidance sheets (see Appendix A) during the peer review for the first topic, most of the peer comments were perfunctory, made only to answer the instructor’s questions on the guidance sheets and principally revolved around grammar. Seeing the students’ misinterpretations the guidance sheets as a series of questions for them to answer and their confusion about how to make comments the instructor/researcher embarked on a training session.

5.3. Peer review training

The peer review training took place during the second and third writing cycles and consisted of two phases—in-class modeling and one-on-one conference after class. The in-class modeling lasted 2 hours for each writing cycle. Each one-on-one reviewer–teacher conference outside class lasted 30 min. In other words, each student received 4 hours of in-class training and 1 hour of reviewer–teacher conference during the whole training period.

5.4. In-class modeling

The modeling was demonstrated when students were about to perform paired peer review on their first drafts of the second and third essays. The instructor first distributed to the students the guidance sheet and a copy of an essay composed by a former student. Then, she used the think-aloud method to demonstrate how to make comments by using a four-step procedure: Clarifying writers’ intentions, identifying the source of problems, explaining the nature of problems, and making specific suggestions (Min, 2005). Techniques for implementing each step were modeled as follows.

While trying to clarify the writer’s intention, the instructor articulated questions like “Do you mean that . . .?” and “Are you saying . . .?” when she was relatively certain of what the writer was trying to convey. When she was uncertain of the writer’s intention, the instructor would model the pointing technique (Stanley, 1992), locating the trouble source (e.g., specific phrases or cohesive gaps) and raising questions such as “What do you mean by . . .?” or “I do not get this” to prompt the writer to explain or revise his/her ideas.

Next, the instructor modeled how to identify problems and explain the nature of problems. If the instructor was certain of the writer’s intention, she would identify the problem and explain why she thought it was problematic. If she had no idea of the writer’s intention, she would refer to the sentences immediately preceding or following the problematic area and articulate what she expected to read given the surrounding contexts. The instructor emphasized to the reviewers that they needed to have logical reasoning to explain why they thought a certain part problematic to convince the writer to accept their comments. Without solid reasoning, even good suggestions are likely to be ignored (Min, 2003).

Finally, the instructor demonstrated how to make suggestions by giving specific examples. Depending on the problems, she would provide a specific definition of a misused phrase and a more appropriate one according to the context, remind the writer to discuss ideas from the same personal perspective, or suggest a specific idea to enrich the content. The instructor informed the students that writers might not adopt their suggestions. Yet, they may have noticed the problems and could work out a solution by themselves.

After the modeling, students were asked to form peer-review dyads on their own, follow the probing questions on the guidance sheets and the four-step procedure, and write and
number their comments on paper according to the order of occurrence of potential problems in the draft. Each student was required to review two different drafts in class and give the written commentary to their partners for revision before the class ended. Following Mittan’s (1989) suggestion, the instructor informed the reviewers that she would grade their commentary to hold them “accountable for their responses” (Ferris & Hedgcock, 2005, p. 233) and to ensure their effort to help each other so that all could progress as writers in class. Writers were allowed 1 week to revise their first drafts at home and were informed of their ownership of the texts. If they did not think their reviewers’ suggestions fit their original ideas, they could disregard them but needed to explain in their revision why they felt certain suggestions failed to work.

5.5. Teacher–reviewer conference

The following week, the instructor collected writers’ drafts, revisions and reviewers’ comments, and checked them carefully. She categorized reviewers’ comments into three groups—check plus, check only, and check minus. Then she scheduled a 30 minutes conference with each reviewer to discuss with them why they received one of the three scores and how to refine their comments so that they were more understandable to the writers. In the one-on-one conference, the instructor offered “content” assistance by working with reviewers on modifying comments that failed to follow the four-step procedure. For example, one reviewer commented on the meager content of a writer’s essay about Yo Yo Ma’s success and suggested adding more examples to enrich the content. The instructor reminded the reviewer to explain why she thought the content meager to make this comment more convincing. She also discussed with the reviewer as to how to make her original comment more specific. The researcher also provided “procedural” assistance to direct reviewers’ attention to the corresponding questions on the guidance sheets through use of various oral prompts (e.g., “Does the third paragraph have a topic sentence that echoes the second main idea of the thesis statement?”).

5.6. Data collection procedure

After the peer review training, the students were required to compose the first draft of the fourth essay at home and bring it to class for peer review. After students completed peer review in class, the instructor allowed writers 1 week to revise their drafts at home. Then she collected their drafts, revisions, and peers’ written feedback of the first and fourth essays the following week for analysis.

Although these two topics appear to be tapping students’ different knowledge bases, they are comparable in that both were expository essays, requiring students to use their analytical ability to either explain the advantages/disadvantages of a certain technological invention or the sources of cultural misunderstandings as well as to synthesize their sources of information in a coherent manner. The average number of words produced for the first and fourth topics were 263 and 257 words, respectively. Two trained independent raters, both of whom have taught EFL writing for more than 8 years, and the instructor carefully read the drafts and revisions prior to and post peer review training to assess essay quality. Then they compared the first drafts, peer feedback, and revisions (a total of 72 texts, 36 texts prior to and post training, respectively) to locate comments that students incorporated into their revisions and to identify the types, levels, and functions of revisions.
5.7. Analysis

Both qualitative and quantitative analyses were performed. Regarding the qualitative analysis, the instructor and two independent raters employed a multiple-trait approach (Hamp-Lyons, 1991a,b) to assessing revision quality prior to and post peer review training. The instructor also conducted interviews with the two independent raters and documented their opinions on types of revision contributing to enhanced text quality. The quantitative analysis, on the other hand, was mainly a text analysis, comparing the number of responders’ comments used in revisions and the ratios of peer-triggered revisions before and after peer review training. Paired t-tests were run to detect any significant mean difference in aforementioned items before and after peer review training. The quantitative analysis also centered on reporting the number of types, sizes, functions of revision, and the types of revision that lead to better text quality. The researcher sequenced the quantitative analysis after the qualitative assessment to avoid the subtle but powerful negative impact of factors not included in multiple-trait measurement (e.g., grammar and spelling) on raters’ judgment of essay quality.

5.8. Qualitative analysis

5.8.1. Criteria for text improvement

Differing measures of text improvement have been employed in different studies. Some considered improved grammar as characteristics of enhanced text quality (Bardovi-Harlig & Bofman, 1989; Celce-Murcia, 1992; Hedgcock & Lefkowitz, 1992). Others deemed appropriateness, sufficiency, and organization of information as much more important criteria in determining text quality (Anson, 1989; Coulthard, 1994; Flower, 1984; Nystrand, 1986; Nystrand, Greene, & Wiemelt, 1993; Sato, 1991). In this study, the two independent raters and the researcher identified improvements on macro features such as idea development, sufficiency, and organization of information as signs of enhanced quality.

Given the previous three criteria, this study employed a “multiple-trait” approach (Ferris & Hedgcock, 2005, p. 315) to assessing student essay quality prior to and post trained peer review. According to Hamp-Lyons (1991b), multiple-trait instruments consider only “the most salient criteria or traits” associated with the writing task, as opposed to a general holistic scoring rubric that evaluates “every element of writing ability that may be manifested in the context” (p. 248). This kind of assessment generates enhanced concurrent and predictive validity given that the more focused trait-based criteria allow raters to resolve differences and reach agreements more easily than with a holistic rubric (Hamp-Lyons, 1991a,b).

Although the three raters did not develop a precise scoring rubric containing detailed descriptors to assign scores to students’ revisions, they agreed to specifically look for the author’s explicit position in the first essay (The advantages/ disadvantages of ... (a new technical invention)), an impartial and accurate presentation of two cultural perspectives in the last essay (An Analysis of an Incident of Cross-Cultural Misunderstanding), and the relevance of directly quoted or indirectly paraphrased information to their positions when evaluating the “appropriateness” of students’ drafts and revisions. Concerning sufficiency, the raters looked for a well-defined thesis and sound synthesized ideas from cited examples and counterexamples. Regarding organization of information, they looked for a direct expression of viewpoint in thesis statements and paragraphs, as opposed to an indirect way preferred by most Chinese writers (Scollon, 1991), logical order, paragraph coherence and transitions, and a restatement of main ideas in conclusion.
5.9. Quantitative analysis

The framework for analyzing types, sizes, and functions of revisions was modeled after Sengupta (1998). She used Falvey’s (1993) taxonomy for corporate revision. ¹

5.9.1. Type of revision

Type of revision is similar to Faigley and Witte’s (1981) surface and text-based change taxonomies. The surface change is further divided into the following subcategories—addition, deletion, substitution, permutation (rephrasing information), distribution, consolidation, and reordering. The text-based changes were further divided into micro-text-based changes and macro-text-based changes. Each contained similar subcategories as those in the surface change category. The difference among surface changes, micro-text-based changes, and macro-text-based changes is that surface changes do not change the overall meaning of the original sentence, whereas micro-text-based changes affect a group of sentences, paragraphs, or the entire text but do not alter the summary of a text. Macro-text-based changes, on the other hand, change the overall summary of the text, altering the direction or the gist of the idea presented. Fig. 2 shows the three main types of revision and their subcategories.

¹ The term “corporate revision” refers to texts taken from “corporations” such as global companies, trading companies, and public institutions, which have to be carefully revised to comply with laws in case there are legal ramifications as a result of misinterpretation or ambiguity (personal communication). The rationale for using this taxonomy for data analysis in this study is that this taxonomy is based on authentic texts which are to be read by the public. In a similar vein, essays composed by students in this writing class were also authentic, in that they were read and responded to by the whole class. As a result, they need to be carefully revised to avoid ambiguity, unexpected interpretation, or plagiarism.
5.9.2. Size of revision

Size of revision refers to the linguistic unit of change, in increasing size: punctuation, word, phrase, clause, sentence, and paragraph.

5.9.3. Function of revision

Function of revision refers to “the analysis based on the textual clues that the researcher can make about the writer’s purpose” (Sengupta, 1998, p. 118). Falvey (1993) identified five functions of revision: grammatical, cosmetic (meaning-preserving changes), texture (make the text more cohesive and coherent), unnecessary expression, and explication (make the text more explicit).

An example of coding students’ writing by type, size and function is as follows:

First draft Nowadays, GM food is developing quickly throughout the world. All scientists are trying to find the way to regenerating the crops to let people live better than before. For example, they use modern biotechnology and put genes from an arctic fish into strawberries to increase the tolerance of cold. It sounds strange, but this kind of GM food may become part of our daily lives. However, those GM foods contain potential danger.

Feedback 1 Be careful of the word better. Is this the exact word that you want to express? Or what you want to express is “comfortable”, “easy”, “pleasant”? If you meant “better”, then be careful of the following sentence where you contradicted yourself by saying “they contain potential danger.”

Feedback 2 Examples should not be put in the introduction paragraph. You may want to explain it in the body paragraph.

Second draft Nowadays, GM food is developing quickly throughout the world. All scientists are trying to develop the standard of living. The techniques of modern genetics can be used in many ways to control the variety of plants, thus likely raising the possibility of health hazard.

The writer followed her first reviewer’s suggestion and corrected the logic in the introduction by deleting conflicting words and sentences and regenerated different ideas. The raters coded this revision as both macro-text-based deletion and addition for types of revision because the deletion of the last two sentences of the first draft and addition of the last sentence of the second draft changed the thesis of the entire essay from addressing both merits and shortcomings of GM foods to discussing their potential dangers to human beings. The deletion of conflicting sentences was an apparent result of peer feedback, but the inspiration for addressing positive effects of GM foods was probably either a result of peer negotiation or self-correction. The size of revision was coded paragraph because the writer revised more than one sentence. The function of revision was coded explication because the writer clearly indicated her intention in a thesis statement, informing the reader of her intention to discuss the possibility of modern genetics jeopardizing health.

In the meantime, the writer also took the second reviewer’s advice by moving the “arctic fish gene and strawberries” example to the third paragraph to support her argument that such gene transference would cause potential health hazards due to the unexpected toxins and allergens that might have been inadvertently created during the process. All raters agreed that this revision was micro-text-based reordering for types of revision. For size of revision, this revision was at the sentence level. In terms of function, this revision was coded “texture” because it made the text more coherent (the example substantiated the argument in the third paragraph). The author changed her thesis from covering both advantages and disadvantages in the first draft to addressing only shortcomings in the revision. Her revision reflected a much clearer stance and more coherent writing. For more detailed explanation and coding of the taxonomy, see Appendices B–D.
5.10. **Raters’ calibration session**

5.10.1. **Assessing revision quality**

The two independent raters and the instructor/researcher attended a calibration session during which they assessed three randomly chosen pairs of draft and revision for both first and fourth essays with one of the results—original better, revision better, or no change. If there was a difference in opinion, the result would be the one shared by two of the three raters. Then, the three would exchange opinions on how they rated the draft and revision by the criteria they reached to further calibrate their assessment standards. After the “norming” session, the three proceeded to rate the texts independently. There was no case in which each rater had a different opinion on a single set.

5.10.2. **Categorizing types, size, and functions of revision**

A week after the raters assessed the text quality, they gathered again to code another three pairs of draft and revision for both first and fourth essays. The inter-rater reliability was 97%, 99%, and 89% for types of revision, size of revision, and function of revision, respectively. They did not encounter difficulty when classifying revision types and sizes but had different opinions on categorizing functions such as cosmetic and texture. They agreed to define cosmetic changes as meaning-preserving changes, and texture changes as cohesion within sentences (e.g., subject-verb agreement) and coherence among sentences (e.g., consistent personal perspective). Then the three raters proceeded to code the remaining texts. The inter rater reliability reached 99, 99, and 96% this time.

5.11. **Findings amount of written feedback incorporated into revision**

To answer the first question regarding the ratio of peer feedback in writers’ revisions prior to and post peer review training, the researcher first tallied the total number of reviewers’ comments and that of reviewers’ comments writers responded to in revision both prior to and post peer review training. Then she calculated the percentages by dividing the total of reviewers’ comments to which writers responded by the total of reviewers’ comments. Results showed that prior to peer review training reviewers generated 130 comments, only 42% \((n = 54)\) of which were incorporated by writers in revision. In contrast, after peer review training, reviewers generated 193 comments, 77% \((n = 149)\) of which were incorporated in revised drafts. Table 1 shows a general picture of the mean differences in the numbers of reviewers’ comments and in the

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.D.</th>
<th>Std. error mean</th>
<th>t</th>
<th>df</th>
<th>Sig. (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feedback 1/feedback 2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>−3.8611&lt;sup&gt;c&lt;/sup&gt;</td>
<td>5.9754</td>
<td>1.4084</td>
<td>−2.741</td>
<td>17</td>
<td>.014&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Used feedback1/used feedback2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>−5.277</td>
<td>2.1366</td>
<td>0.5036</td>
<td>−10.480</td>
<td>17</td>
<td>.000**</td>
</tr>
</tbody>
</table>

<sup>a</sup> Feedback 1 refers to the total number of peer comments before training, whereas feedback 2 refers to that after training.

<sup>b</sup> Used feedback 1 refers to the number of peer comments incorporated into revisions before training, whereas used feedback 2 refers to that after training.

<sup>c</sup> The negative means and t values indicate that the numbers of feedback 1 and used feedback 1 are smaller than those of feedback 2 and used feedback 2.

<sup>*</sup> \(p < 0.05\).

<sup>**</sup> \(p < 0.01\).
numbers of comments writers incorporated in revision before and after training. As clearly demonstrated, the numbers of total comments produced and those incorporated into revision after peer review training were significantly higher than those before training (for the number of comments, $t = -2.741, p < .05$; for the number of incorporated comments, $t = -10.480, p < .01$).

5.12. Proportions of peer-influenced revisions in total revisions

Despite the fact that students incorporated a significantly higher number of peer review comments into their revisions after peer review training, this only shows that trained peer review exerted a significantly greater influence on writers’ revisions than the untrained one. It does not assure that these peer-triggered revisions take up most of the changes writers made in their revisions. The exact strength of this trained peer review feedback on writers’ revisions is still unanswered. In order to better understand the strength of the effect of this peer review training on revision, it is important to examine the extent to which revisions were a result of peer feedback, both prior to and post peer review training.

An examination of the revisions made by the writers revealed that there were 80 revisions prior to peer review training and 165 revisions post training. Of the 80 revisions made before training, 54 were in response to untrained peer feedback. Of the 165 revisions made after training, 149 were in response to trained peer review feedback. The revisions in response to peer feedback were 68% (54/80) before training and 90% (149/165) after training. Table 2 shows that there was a significant difference between the total revisions writers made before and after peer review training ($t = -12.806, p < .05$). There was also a significant difference in the revisions as a result of peer feedback before and after peer review training ($t = -3.778, p < .05$). The latter finding showed that trained peer feedback did have a significantly higher impact on students’ revisions after peer review training.

5.13. Revision quality

Given that 68% of the revisions were made in response to peer comments before peer review training and 90% of them were a result of peer comments after training, one cannot help but wonder about the quality of such revisions. The raters deemed approximately 13% of the revisions prior to training as better, 9% of the revisions inferior to the original, and 78% unchanged. In contrast, they considered approximately 72% of the revisions superior, 19% of the revisions poorer, and 9% unchanged. These results show that trained peer review did enhance the

<table>
<thead>
<tr>
<th>Difference</th>
<th>Mean deviation</th>
<th>Std. Mean</th>
<th>S.E.</th>
<th>$t$</th>
<th>Sig. (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total revision 1/total revision 2</td>
<td>-4.7222&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.5645</td>
<td>0.3687</td>
<td>-12.806</td>
<td>.000&lt;sup&gt;**&lt;/sup&gt;</td>
</tr>
<tr>
<td>RPF1/RPF2</td>
<td>-1.250</td>
<td>1.4037</td>
<td>0.3308</td>
<td>-3.778</td>
<td>.002&lt;sup&gt;**&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>*</sup> $p < 0.05$.
<sup>**</sup> $p < 0.01$.

<sup>a</sup> Total Revision 1 means the total revisions made before peer review training, and total revision 2 means those made after peer review training.

<sup>b</sup> RPF 1 means the proportion of peer-triggered revisions before peer review training, and RPF 2 refers to those after training.

<sup>c</sup> The negative means and t values indicate that the numbers of total revisions and peer-triggered revisions before peer review training were smaller than those after peer review training.
quality of students’ revisions, given that 90% of their revisions were based on trained peer review feedback. Table 3 indicates how the three independent raters evaluated the texts produced before and after peer review training. Most of the first drafts prior to and post training, according to the raters, lacked explicit positions, developed ideas, logic, and good organization. The revisions produced prior to peer review training, similar to those found in Sengupta’s (1998) study, were mainly at the word level that did not effect any “sweeping alterations” of propositions (p. 121). By contrast, most of the revisions post peer review training were improved in terms of idea development, unity, and organization. Thus, the overall quality was enhanced.

5.14. Type of changes in revisions

5.14.1. Revision types

The fourth research question concerns the kinds of revisions writers made in response to peer feedback in revision. A tally of revision types revealed that substitution (20%), permutation (19%), and reordering (18%), especially at the micro-text-based level, ranked as the highest three. These findings partially corroborated previous studies that have investigated revision (Sato, 1991; Sengupta, 1998). Both Sato’s Japanese college students and Sengupta’s Hong Kong secondary students also used substitution most frequently, but at the surface level. Table 4 shows the percentages of the types of feedback triggering revisions for the student writers in this study.

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Frequency of types of revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microstructure</td>
<td>Surface changes</td>
</tr>
<tr>
<td></td>
<td>Macrostructure changes</td>
</tr>
<tr>
<td>Addition</td>
<td>6 (4%)</td>
</tr>
<tr>
<td>Deletion</td>
<td>3 (2%)</td>
</tr>
<tr>
<td>Substitution</td>
<td>9 (6%)</td>
</tr>
<tr>
<td>Permutation</td>
<td>12 (8%)</td>
</tr>
<tr>
<td>Distribution</td>
<td>4 (3%)</td>
</tr>
<tr>
<td>Consolidation</td>
<td>3 (2%)</td>
</tr>
<tr>
<td>Reordering</td>
<td>8 (5%)</td>
</tr>
<tr>
<td>Total</td>
<td>45 (30%)</td>
</tr>
</tbody>
</table>

a The percentages do not add up to exact 100% due to rounding.
5.15. Revision sizes

With regard to size of revisions, the most frequent revision occurred at the level of sentence (32%), closely followed by paragraph (20%) and word (20%). The least revised part in terms of size of revisions was at the level of symbol. Only seven instances were recorded. Sengupta (1998) also found the sentence to be the unit that received most revisions in her study. Table 5 shows the revisions across different sizes of linguistic units.

5.16. Revision functions

The most common function of revision for this study was texture revision (39%), rendering the text more cohesive and coherent. Explicature (29%) and cosmetic (21%) were the second and third concerns (see Table 6). It is likely that texture (concerning coherence) and explicature (concerning explanation) were the most commonly perceived functions of revision because two of the principal foci of the guidance sheet used in peer review training were format and content. The reviewers had been constantly reminded in the guidance sheets to look for thesis statements, topic sentences, transitions among paragraphs, and the clarity of content. As a result, reviewers were especially demanding about the whereabouts of those signal sentences and transitional conjunctions. In addition, reviewers were also trained to clarify writers’ intentions before attempting any suggestions. It is unsurprising that most reviewers requested writers to further explain their intentions, to which writers complied by adding more content.

Interestingly, grammatical change (4%) was less common among writers in this study. A closer look at their first drafts revealed a lot of grammatical mistakes. Yet, their reviewers appeared to ignore those errors as long as they did not obscure the intended meaning. Such a keen attitude toward understanding the content is probably an effect of the training. Before peer review training, most reviewers, perhaps due to their former language training

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Revisions across different linguistic units</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Size</strong></td>
<td><strong>Number of revisions</strong></td>
</tr>
<tr>
<td>Symbol</td>
<td>3</td>
</tr>
<tr>
<td>Word</td>
<td>30</td>
</tr>
<tr>
<td>Phrase</td>
<td>21</td>
</tr>
<tr>
<td>Clause</td>
<td>17</td>
</tr>
<tr>
<td>Sentence</td>
<td>48</td>
</tr>
<tr>
<td>Paragraph</td>
<td>30</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>149</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Functions of revision</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Function</strong></td>
<td><strong>Number of revisions</strong></td>
</tr>
<tr>
<td>Grammatical</td>
<td>6</td>
</tr>
<tr>
<td>Cosmetic</td>
<td>31</td>
</tr>
<tr>
<td>Texture</td>
<td>58</td>
</tr>
<tr>
<td>Unnecessary</td>
<td>11</td>
</tr>
<tr>
<td>Explicature</td>
<td>43</td>
</tr>
</tbody>
</table>
background and misunderstanding of the purpose of peer review, tended to correct grammatical errors only. During the peer review training, the instructor requested the reviewers to attend to writers’ intentions and assist them in expressing their ideas in a comprehensible fashion. It is possible that reviewers thus complied with the instructor’s request and attended to meaning more than they did to grammar during peer review, as evidenced by few grammatical revisions (4%) and many explanations (29%), which were to render writers’ intentions clearer to reviewers.

Sengupta (1998) also found a low percentage of grammatical revision, but suggested a different factor contributing to the low ratio of grammatical revision among his subjects: learners’ limited linguistic resources. She expressed reservations about those secondary school students’ English proficiency and suspected that most were not proficient enough to pass the linguistic threshold to make grammatical revisions. Writers in this study, on the contrary, were second-year English majors whose English proficiency was presumably higher than that of those in Sengupta’s (1998) study. Low language proficiency was less likely a cause of low grammatical revision. Given that students under focus were required to deliberately hold back their urge to correct grammar errors and to center their attention on meaning first, instructional intervention was more likely the main cause of low grammatical revisions found in this study.

To summarize, substitution and permutation at the micro-text-based level were the most common revision types, immediately followed by reordering at the same level. Corroborating this finding is that revisions occurred most frequently at the sentence level, closely followed by paragraph and word level revisions. The main functions of word substitution and rearrangement at the sentence and paragraph levels are texture (coherence), followed by explicature.

5.17. Revisions leading to better texts

The kinds of revision that led to better texts in this study were substitution, permutation, and reordering at the text-based level. As discussed earlier, the three types of revisions occurring most frequently were substitution (20%), permutation (19%), and reordering (18%). Substitution and permutation are surface changes whereas re-ordering is a text-based change at the micro- or macro-level. How did the three different types of revisions interact with one another to achieve quality texts?

An examination of the revised drafts revealed that most revisions in this study were centering on reordering information to make it fit more appropriately into the format of an essay (e.g., the placement of a thesis statement and a topic sentence), rendering it easier for reviewers to comprehend their intentions. Sometimes paragraphs were rearranged or combined to emphasize the importance of main ideas across several paragraphs to achieve coherence. Once ideas became clear and organized through re-ordering of sentences and paragraphs, writers turned to polishing their register by rearranging phrases (permutation) or substituting informal and inexact words with more formal and precise ones to enhance its professional look.

The frequency distribution of revision size also corroborated this interpretation, given the most frequent revision occurring at the level of sentence (32%), closely followed by paragraph (20%) and word (20%). Sentence- and paragraph-level re-ordering reflected micro- or macro-level text-based revisions, and substitution and permutation at the word level indicated cosmetic meaning-preserving changes. Finally, the functions of revision also supported this interpretation because the most common function of revision was texture revision (39%). Most revisions were geared toward rendering the text more coherent by rearranging sentences and paragraphs to maintain the focus on the same idea or in the same perspective, usually the third person.
More appropriate phrases were used to explain the writers’ intentions (explicature, 29%) and more precise words to increase the register of the writing (cosmetic, 21%).

The previous interpretation was also triangulated with interview results. The raters agreed that re-ordering presented the information in a more organized fashion, satisfying their expectation and letting them know immediately the writers’ intentions. “Some of the first drafts lacked a thesis statement. Others lacked main ideas. But the writers were able to remedy those situations by providing a sentence or moving a sentence in the main body to the place of the thesis statement in revision. This revision gave me a much clearer picture of the writers’ intentions.” They also agreed that most of the substitutions and permutations, if considered separately from micro- or macro-level changes such as re-ordering, consolidation and distribution, did not change much of the meaning of the original compositions. As succinctly pointed out by one rater, “Some revised texts were mainly substitutions of words and phrases, which had little to do with idea development, clarification, or organization. Had these surface-level substitutions and permutations not co-occurred with other micro- and macro-structure text-based revisions such as reordering, consolidation, and distribution, those superficial changes would not have had as much impact as they did on enhancing overall text quality.” As clearly shown in the previous interviews, the raters deemed micro- and macro-level text-based revisions (re-ordering, consolidation, and distribution) major causes of improvements because such revisions enhanced sufficiency, relevance, and organization of information, three criteria for quality texts. This perspective is also shared by Paulus (1999) who reported a moderate positive correlation between the percentage of macrostructure changes and the amount of overall essay improvement.

6. Discussion

The finding that trained peer review has a positive impact on refining reviewers’ comments and communication strategies and writers’ subsequent revisions is not new, given similar findings reported in experimental (Berg, 1999) and classroom studies (Paulus, 1999; Stanley, 1992; Zhu, 1995). Nevertheless, this study has been the first attempt to delineate the exact effect of trained peer review on student revision. Unlike Berg’s (1999) study, which did not seek to disentangle the effects of trained peer response and self-feedback on revision quality, the researcher was able to trace the impact of the trained peer feedback by using written commentary. Given that each comment was produced on a piece of paper numbered by reviewers according to the order of occurrence of potential problems denoted by the same number in writers’ first drafts it is easy to distinguish revisions made in response to peer feedback from those that did not. In this regard, peer written feedback appears to be a more fruitful approach to grappling with the exact relationships between peer comments and revision strategies and writing outcomes than peer (oral) response groups. Such an approach is also utilized to examine the effects of teachers’ written commentary on students’ revisions (Conrad & Goldstein, 1999).

The result demonstrated that 77% of the trained peer review feedback was incorporated into students’ revisions, which constituted 90% of the total revisions. This high percentage of peer feedback incorporation is in contrast to that before these students received peer review training (39%) and those reported in the literature (Chou, 1999; Connor & Asenavage, 1994; Mendonca & Johnson, 1994; Paulus, 1999; Tang & Tithecott, 1999; Tsui & Ng, 2000). Such a high rate may suggest two things: that student writers in this study found trained peer feedback helpful and were willing to use it in their subsequent revisions; or that they did not find trained peer review feedback helpful but used it nonetheless, perhaps to avoid taking further trouble to explain to the instructor in writing why they did not employ it. The latter interpretation appears less plausible.
since previous interviews with the same group of writers revealed that most of them (15 out of 18) found the trained peer review feedback helpful, particularly in focusing their ideas and enriching the content by viewing things from different perspectives. Some even attributed their revision improvement to their “classmates’ helpful feedback” (Min, 2005, p. 302). This finding should bolster those ESL/EFL writing instructors who have integrated peer review into their writing classes and those who were disheartened by Nelson and Carson’s (1998) claims that students do not find their classmates’ feedback particularly helpful.

Two important factors may shed light on the success of the peer review training in the current study—the individual teacher–reviewer conferences and the instructor’s grading peer review comments. Previous research suggests that classroom demonstration alone is not adequate for successful implementation of peer review skills. Peer response training involved with video or teacher demonstration in class does not ensure a high rate of incorporating peer feedback into revision (Connor & Asenavage, 1994) or a positive training effect on shaping revision quality (Tang & Tithecott, 1999) because students lack opportunities to transform their declarative knowledge into procedural knowledge and to consult teachers when problems arise in class. Additional assistance outside of the classroom is needed. Connor and Asenavage’s (1994) specifically recommended that “more extensive and specific peer response training with follow-up [italics mine] should be implemented” (p. 267). The inclusion of after-class teacher–reviewer conferences in addition to in-class modeling during peer review training in this study is a direct response to their calls. An analysis of the journal entries of the same group of reviewers (Min, 2005) revealed that some would not have known how to perform peer review or refine their review comments if there had not been teacher–reviewer conferences to address their individual difficulties, suggesting that these conferences helped enhance the quality of peer review comments and thus the effectiveness of peer review training.

Another instructional technique – grade assignment to peer review comments – may also have enhanced the effect of the peer review training. The instructor’s grading appeared to serve as a strong incentive for this group of reviewers to invest time and effort in following the four-step procedure while offering feedback, although the original intention was to hold them accountable for their written comments. Reviewers showed keen interests in how to refine their written commentary to obtain a higher grade during teacher-reviewer conferences all the time, which perhaps contributed to quality feedback. Future studies and classroom instructors may conduct interviews to obtain reviewers’ viewpoints on the impact of such an instructional technique on their written feedback. More research is also needed to compare the quality of peer review comments receiving teachers’ evaluation with those without it.

Similar to previous research investigating the impact of peer feedback on revision types and quality (Berg, 1999; Connor & Asenavage, 1994; Paulus, 1999), the current study adopted Faigley and Witte’s (1981) taxonomy to categorize students’ revisions. Although this analytical model deals with the number of revision types, it suffers two limitations (Ferris, 2003). First, this simple counting taxonomy fails to put different numbers of revision types into perspective; it is thus likely to mislead researchers to stereotyping students as form-focused revisers who usually make more meaning-preserving changes than macrostructure ones. As argued by Ferris (2003), the common length of student essays (between 500 and 1000 words) may warrant only “a few text-based, macrostructure changes” but necessitate “dozens or even hundreds of possible surface formal changes” (p. 36). Given this argument, it is unsurprising that researchers almost always document more surface changes than text-based ones in students’ revisions. Yet, this difference in the numbers of revision types should not be
automatically translated as students’ inability to revise or possible failure of peer response/review training. Researchers should keep the differential numbers of revision types in perspective when interpreting their significance.

The second limitation of Faigley and Witte’s (1981) categorization scheme, according to Ferris (2003), lies in the fact that it does not address the possible effects of revision types on revision quality. In actuality, researchers attempting to explore the relationship between revision types and text quality have reported that sentence level, meaning-preserving revisions did not necessarily reflect quality changes (Paulus, 1999). Indeed, the effectiveness of a writing sample is much greater than the sum of its parts. Linking the quality of a text to revisions at the sentence level decreases the interconnectedness of written discourse and gives the false impression that writing can be understood and fairly assessed by analyzing isolated text features (Hillocks, 1995; White, 1994). “Only meaning-level changes” at the macrostructure level may be “the ones that improved the essays” most significantly (Paulus, 1999, p. 282). Interviews with the two independent raters in this study also corroborated this finding that changes at the intersentential level are most likely to enhance text quality.

Given that Faigley and Witte’s (1981) revision taxonomy does not concern itself with overall revision quality, and that no empirical evidence shows a strong positive relationship between the amount of revisions and revision quality (Paulus, 1999), the issue of finding an appropriate approach to writing assessment other than tallying accumulated revisions becomes crucial. Unlike previous researchers who employed generic holistic (Berg, 1999) or analytic scoring (Paulus, 1999) to evaluate revision quality, this study adopted a multiple-trait approach which focused raters’ attention mainly on macro features such as idea development, sufficiency, and organization. Utilizing this approach, raters in this study could avoid assigning the same score to two texts with entirely distinct sets of characteristics either through impressionistic judgment (as in holistic rating) or through arbitrary equal weighting to macro and micro features (as in analytic scoring). This multiple-trait approach provides higher content and construct validity and yielded enhanced concurrent and predictive validity (Hamp-Lyons, 1991a,b). Given its inherent nature in increasing concurrent validity, it is unsurprising that a high inter rater reliability in assessing student essay quality was obtained in this study. Future researchers interested in writing assessment can use this type of measure with a scoring rubric and more refined descriptors to increase the reliability and validity of their assessment scores.

7. Conclusion

Berg (1999) gave a cogent argument for the irreplaceable position of peer review training in ESL/EFL writing instruction, stressing its importance of alerting inexperienced writers to the incongruity between their intended meaning and the actual meaning perceived by their readers. In actuality, inexperienced writers not only lack the ability to detect the dissonance between their intended meanings and readers’ understood meanings (Nold, 1981; Sommers, 1980) but also the prerequisite linguistic resources for revision strategies once such disparity is discerned (Sengupta, 1998). Revision instruction cannot effectively deal with the abovementioned problems among inexperienced writers because it requires them to view their texts from an outsider’s perspective and think of revision strategies alone. The findings of this research demonstrated that without peers’ assistance in both ideas and language, it is very difficult, if not impossible, for them to grapple with these two issues single-handedly. Inexperienced writers must be provided with an avenue to recognizing the mismatched meanings between writers and readers and to addressing these problems. It is only through a step-by-step peer review training
procedure like the one employed in this study that students can be helped to view texts from multiple perspectives and clarify misunderstandings if needed.

The results of this preliminary study cannot be generalized to other contexts due to a lack of control group and the small number of students involved. Researchers can test the strength of combining in-class demonstration and the conference approach after class in experimental studies, but they need to grapple with the ethical issue of deliberately depriving student writers of precious peer review learning opportunities for an extended period of time. Although this study provided an avenue to tracing the impact of trained written peer review on revision, the findings also showed that some of the revisions were not caused by peer comments. The sources of revisions might be others’ or the writers’ own ideas. Future research can involve retrospective interviews with student writers to examine their own revisions and discuss the effects of peer review comments on their decision making (Ferris, 2003).

Acknowledgement

The author would like to thank two anonymous reviewers and the editors for their helpful suggestions during the preparation of this article. This research is partially funded by National Science Council in Taiwan (NSC 91-2815-C-006-088-H).

References


Appendix A

A.1. Guidance sheet for reviewing Multiple-paragraph essays

1. Read the introductory paragraph. Is there a thesis statement toward the end of the introduction? Does the thesis statement contain main ideas? How many main ideas are there? Please underline the thesis statement and mark 1, 2, or 3 on each main idea. Are these main ideas at the same level of generality? Are they sequenced in accordance with importance? If you cannot find a thesis statement, drawing on what you have read so far, what do you expect to read in the following paragraphs? Summarize it in one sentence and show it to your partner.

2. Now read the first few sentences in the second paragraph. 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 thesis statement to make sure that you understand the main ideas. Did the author talk about the first main idea in the thesis statement? If not, remind him/her that he/she should. Are there any concrete examples or explanation in this paragraph to support the main idea? Are they well balanced (in terms of sentence length and depth of discussion)? Are they relevant and sequenced properly? Is there any direct quotation or paraphrased information in this paragraph? Is the quotation supporting the argument the writer has made? Check the original source if your partner wrote a paraphrase to make sure that the paraphrase reflects accurate information.

3. Read the first sentence of the third paragraph. Did your partner use any transitions to connect this paragraph with the previous one? If not, can you suggest one? Is there a topic sentence that corresponds to the second main idea in the thesis statement? Make a suggestion if there is not. Are there any concrete examples or explanation in this paragraph to support the main idea of this paragraph? Are they well balanced (in terms of sentence length and depth of discussion)? Are they relevant and sequenced properly? Is there any direct quotation or paraphrased information in this paragraph? Is the quotation supporting the argument the writer has made? Check the original source if your partner wrote a paraphrase to make sure that the paraphrase reflects accurate information.

4. Read the first sentence of the fourth paragraph. Does this paragraph connect well to the previous one? If not, can you suggest a sentence connector? Is there a topic sentence that corresponds to the third main idea in the thesis statement? Make a suggestion if there is not. Are there any concrete examples or explanation in this paragraph to support the main idea of this paragraph? Are they relevant and sequenced properly? Did your partner use pronouns and paraphrase to avoid repetition? Is there any direct quotation or paraphrased information in this paragraph? Is the quotation supporting the argument the writer has made? Check the original source if your partner wrote a paraphrase to make sure that the paraphrase reflects accurate information.

5. Read the conclusion. Does it begin with a restatement (but different wording) of the thesis statement? If not, suggest one. Does the conclusion move to more general statements on the
topic as a whole? Does the conclusion contain too much irrelevant information to the thesis statement? If yes, make a suggestion.

6. What did you learn from reading this essay, either in language use or content? Is there anything nice you want to say about this essay? Are there any grammatical errors or inappropriate word usage?

Appendix B

B.1. Types of revision

<table>
<thead>
<tr>
<th>Type</th>
<th>Example (changes in boldface)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addition: reviser adds information</td>
<td>First draft: GM foods can increase harvest. Feedback: There seems to be a logic problem. GM foods do not help increase harvest. It is the use of genetic engineering technique that leads to an increase on harvest. Second draft: Planting GM foods can help farmers increase harvest.</td>
</tr>
<tr>
<td>Deletion: reviser deletes information</td>
<td>First draft: In today’s society, GM food is becoming increasingly trendy. That is GM food is a must. Feedback: A trendy thing does not mean that it is a must. Second draft: In today’s society, GM food is becoming increasingly trendy.</td>
</tr>
<tr>
<td>Substitution: reviser substitutes information</td>
<td>First draft: Today, GM food has tremendously benefited farmers around the world,… Feedback: I suggest that you change farmers into people. Second draft: Today, GM food has tremendously benefited people around the world,…</td>
</tr>
<tr>
<td>Permutation: reviser rephrases information</td>
<td>First draft: Cell phones are not just chic gadgets, but with them we can talk to anyone on the planet from just about everywhere. Feedback: Can you use a noun phrase to make the structure more parallel. For example, “…but communication devices that can…” Second draft: Cell phones are not just chic gadgets, but communication devices that can bring people together.</td>
</tr>
<tr>
<td>Distribution: reviser re-writes same information in larger chunks</td>
<td>First draft: Fifty years ago, people were threatened by poverty and starvation. My grandma described it a world riddled with suffering and numerous people were as poor as church mice. Feedback: It looks like the starvation is because of poverty, not lacking of food. Second draft: Fifty years ago, people were threatened by poverty and starvation. My grandma described it a world riddled with suffering. Indeed, without harvests, lots of people were as poor as church mice.</td>
</tr>
<tr>
<td>Consolidation: reviser puts separate information together</td>
<td>First draft: Cell phones with attachable cameras or cameras embedded in them have become ubiquitous that they might be a potential for intruding people’s privacy. Gym lockers, for example, where photography is greatly discouraged since long time ago. Feedback: The second sentence is not a full sentence. You might want to combine it with the first one. Second draft: The ubiquity of cell phones with attachable or built-in cameras might be intruding people’s privacy, especially in private places such as gym lockers, where photography has long been discouraged.</td>
</tr>
<tr>
<td>Re-order: reviser moves information</td>
<td>First draft: GM Food nearly can be seen everywhere you can reach. Feedback: “Nearly can be seen” means people do not see it. They “nearly” see it. Second draft: GM Food can be seen nearly everywhere you can reach.</td>
</tr>
</tbody>
</table>
C.1. Size of revision

<table>
<thead>
<tr>
<th>Size</th>
<th>Example (Changes in bold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol etc.</td>
<td>First draft: Nowadays there have been many farmers growing GM plants. Feedback: You do not start a line with a period. Second draft: Nowadays there have been many farmers growing GM plants.</td>
</tr>
<tr>
<td>Word</td>
<td>First draft: Through the networks of tower transmitters and even satellites, cell phones can help us talk to anyone from anywhere in anytime. Feedback: It is “at” anytime, not “in” anytime. Second draft: Through the networks of tower transmitters and even satellites, cell phones can help us talk to anyone from anywhere at anytime.</td>
</tr>
<tr>
<td>Phrase</td>
<td>First draft: The incredible gadgets, cell phones enrich our lives so much and undoubtedly they are necessaries machines for most people nowadays. Feedback: Do you mean indispensable devices? The size of a machine seems to be larger than that of a cell phone. Second draft: The incredible gadgets, cell phones enrich our lives so much and undoubtedly they are indispensable devices for most people nowadays.</td>
</tr>
<tr>
<td>Clause</td>
<td>First draft: Those people who feel its danger believe that cell phones will do harm to our health. Feedback: It reads like Chinese English. You can say people who are aware of its danger or who feel that it is dangerous. Second draft: Those people who are aware of its danger believe that cell phones will do harm to our health.</td>
</tr>
<tr>
<td>Sentence</td>
<td>First draft: A great deal of research proved that transgenic crops will transfer the characteristics to their relatives. If genes that have the function to resist insect attack are introduced into plants, the establishment of resistant populations of pests will increase soon. And that will cause the loss of biology balance. Feedback: You need to explain why it will cause biological imbalance. First draft: A great deal of research proved that transgenic crops will transfer the characteristics to their relatives, such as the ability of herbicide or insects-resistance. Then the insects will also grow stronger and be able to eat that kind of plants. As a result, the population of pests will increase soon. And under the circle, it will cause the loss of biology balance.</td>
</tr>
<tr>
<td>Paragraph</td>
<td>First draft: Professor Chung-Shau Gim said that some people like to put their cell phones in the pockets of their shirts or pants and this might have bad effects to their hearts and organs in abdomens. He also stated that there were 500 medical research studies pointing out that if human bodies are exposed to EMF, it will decrease our immunocompetence and make the medicines for mastocarcinoma lose its efficacy. Feedback: In the third paragraph, you can add some examples to the 3rd paragraph about how cell phones damage our health. Second draft: Chung-Shau Gim, an associate professor in the physiology research institute of National Defense Medical College, said that, according to 500 medical reports, our immunocompetence would be decreasing and the medicines for mastocarcinoma would lose its efficacy if we are exposed EFFs for a long time. ES, electromagnetic sensitivity syndrome, is a kind of disorder with neurological and allergic-type symptoms. When we are exposed to a new EMF, such as a new computer or new cell phone, we might get ES. The symptoms of ES include headache, eye irritation, dizziness, nausea, skin rash, fatigue, and etc.</td>
</tr>
</tbody>
</table>
Appendix D

D.1. Function of revision

<table>
<thead>
<tr>
<th>Function</th>
<th>Example (Changes in bold)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grammatical:</strong> to make the text grammatically correct</td>
<td>First draft: Through the networks of tower transmitters and even satellites, cell phones can help us talk to anyone from anywhere in anytime. Feedback: It is “at” anytime, not “in” anytime. Second draft: Through the networks of tower transmitters and even satellites, cell phones can help us talk to anyone from anywhere at anytime.</td>
</tr>
<tr>
<td><strong>Cosmetic:</strong> a change which makes the text look better</td>
<td>First draft: The incredible gadgets, cell phones enrich our lives so much and undoubtedly they are necessary machines for most people nowadays. Feedback: Do you mean indispensable devices? The size of a machine seems to be larger than that of a cell phone. Second draft: The incredible gadgets, cell phones enrich our lives so much and undoubtedly they are indispensable devices for most people nowadays.</td>
</tr>
<tr>
<td><strong>Texture:</strong> To make the text more cohesive and coherent</td>
<td>First draft: Those people who are aware of its belief that cell phones will do harm to our health Feedback: You use “those people” at the beginning of this sentence but “our” at the end. You might want to stick to the same personal perspective by using “their”. Second draft: Those people who are aware of its danger believe that cell phones will do harm to their health.</td>
</tr>
<tr>
<td><strong>Unnecessary expression:</strong> to take away unnecessary information</td>
<td>First draft: Mary comes from America, and of course, she has no idea about Chinese lucky money. Feedback: The use “of course” is unnecessary. Second draft: Mary comes from America, and has no idea about Chinese lucky money.</td>
</tr>
<tr>
<td><strong>Explicature:</strong> to make the information in the text more explicit</td>
<td>First draft: A great deal of research proved that transgenic crops will transfer the characteristics to their relatives. If genes that have the function to resist insect attack are introduced into plants, the establishment of resistant populations of pests will increase soon. And that will cause the loss of biology balance. Feedback: You need to explain why it will cause biological imbalance. Second draft: A great deal of research proved that transgenic crops will transfer the characteristics to their relatives, such as the ability of herbicide or insects-resistance. Then the insects will also grow stronger and be able to eat that kind of plants. As a result, the population of pests will increase soon. And under the circle, it will cause the loss of biology balance.</td>
</tr>
</tbody>
</table>

Hui-Tzu Min is an associate professor currently teaching in the Department of Foreign Languages and Literature at National Cheng Kung University, Taiwan. Her areas of interest include children’s and adults’ second language acquisition and ESL/EFL literacy.