英語教學 *English Teaching & Learning*
39.1 (Spring 2015): 1-28
DOI: 10.6330/ETL.2015.39.1.01

Motivating Revision by Integrating Feedback Into Revision Instructions

Shu-Chen Huang[a]

Abstract

Second language writing teachers usually provide individual written feedback on learner drafts. Research shows that such feedback often fails to improve learner texts partly due to communication problems in the way feedback is given. Although directly teaching revision may help learners enhance the quality of their writing, many such revision instructions are an independent component added to traditional feedback. It is of interest to determine if feedback to learner drafts could be incorporated into revision instructions and help learners improve their texts. In an EFL course of 38 freshmen, the researcher implemented feedback principles in designing three revision lessons. Teaching points were based on common problems identified. After three rounds of writing, each consisting of a draft-instruction-revision sequence, the drafts and revisions were evaluated by two independent raters. In addition, learner experience was examined through an open-ended questionnaire. Findings indicate that learners improved from each draft to its revision, and the effect was more obvious at global/textual than at local/linguistic levels. However, it is inconclusive whether this positive effect continued from one task to the next. Although learner response to the feedback-embedded revision instruction was generally positive, the participating students also indicated a desire for traditional individual feedback.

Key Words: written feedback, revision instruction, assessment for learning

[a] Foreign Language Center, National Chengchi University. E-mail: huang91@nccu.edu.tw

INTRODUCTION

Providing individual feedback to learners on their writing assignments is an important part of second language writing instruction (Ferris, 2003). There have been abundant discussions in the literature on how teacher feedback could be made more effective (e.g., Chandler, 2003; Ellis, Sheen, Murakami, & Takashima, 2008; Ferris, 2003; Lee, 2007; Montgomery & Baker, 2007). Among them, some provide evidence for the efficacy of written corrective feedback on linguistic accuracy (e.g., Bitchener & Knoch, 2010; Ferris, 2006; Polio, Fleck, & Leder, 1998); others deal more with learner improvement beyond error correction (Butler & Britt, 2011; Myhill & Jones, 2007).

Despite the possible benefits, it has been attested that some students pay little attention to teacher feedback or do not understand how to act upon it (Sadler, 2010; Wingate, 2010). Reasons for students' non-engagement with feedback include teachers' failure to offer opportunities or strategies for using feedback (Silver & Lee, 2007), and students' low motivation or low self-perception as writers, both of which are usually neglected in second language acquisition studies (Ellis, 2010). It is also suggested that teachers and researchers pay close attention to the ways in which feedback content is communicated, as well as to learners' affective factors. In so doing, teachers may increase the likelihood that students utilize the feedback. In fact, it is generally acknowledged that students who read and acted on the feedback suggestions improved in the areas previously commented upon (Sadler, 2010; Wingate, 2010).

Providing students with individualized written feedback is not the only way to help learners improve the quality of their writing. Some writing teachers do so by explicitly teaching revision strategies. Two recent studies have sought to deliberately teach revision to the class as a whole. Sengupta (2000), in addition to providing written feedback on individual submissions, taught Hong Kong secondary students revision strategies following the completion of their first drafts. Writing performance was then holistically measured and compared with students who did not learn these revision strategies. The findings demonstrated that explicit teaching of revision strategies, when combined with individual written feedback, had a significant effect on writing scores, especially in facilitating an awareness of discourse-related features in L2 writing. Another study with similar objectives was situated in a first language tertiary context in the U.S. Butler and Britt (2011) found their students underprepared for academic writing, such that they could not write well-structured arguments. In response,

the researchers designed two writing tutorials to help students revise their argumentative essays—an argument tutorial and a global revision tutorial. Student groups completed either or both tutorials independently. It was shown that, regardless of whether they learned argument, revision, or both, all improved their revised submissions, and the improvement level among these groups was indistinguishable.

Sengupta's (2000) study showed that giving individual feedback and teaching revision strategies improved the EFL writing scores of Hong Kong secondary students. On the other hand, Butler and Britt's (2011) study demonstrated that college-level L1 writers improved their writing by learning from written tutorials. It is not clear whether EFL learners could also benefit from revision instructions alone, if individual feedback is absent. In fact, Sengupta's dual approach may sound more ideal than practical for a busy teacher, especially in large classes. As can be seen in recent feedback studies (Bailey & Garner, 2010; Lee, 2011b; Price, Handley, Millar, & O'Donovan, 2010), how teachers' time and effort could be more strategically managed has become a major concern.

As discussed above, with the same objective of having learners improve their drafts, a writing teacher has two approaches at her disposal—the more traditional feedback comments given to individual students, or instructions on revision strategies prepared for the entire student group. The former approach may be more successful in addressing individual learner needs, but it usually suffers from inadequate communication as it mostly takes a peripheral status outside of class. The latter, on the other hand, could be communicated more fully in a face-to-face classroom setting but does not seem to address individual learner problems directly. Although more experienced teachers may be able to skillfully merge these two approaches in their day-to-day lectures, how to consciously plan and carry out such instruction is seldom discussed and may be difficult for less experienced teachers. How can the benefits of feedback and lectures be brought together so that particular learner problems are attended to and, at the same time, communications become more comprehensive and effective? In fact, Hattie and Timperley (2007), in their much-cited review article on the power of feedback, argue that teaching and giving feedback, although usually marking the beginning and endpoint of an instructional program respectively on two opposite ends of a continuum, could and should be brought together. Indeed, instruction should be more effective after a teacher has first assessed where learners are and what they need. Moreover, for feedback to be understood and acted upon by learners, the teacher has to carefully plan the *what, why*, and *how* of the complicated feedback messages.

To answer these questions, the author of this paper, inspired by recent literature on feedback and assessment for learning, attempted to integrate feedback into revision instructions and bring it from the peripheral to the center stage. Unlike those revision tutorials in Sengupta's (2000) and Butler and Britt's (2011), which were independent from teacher assessment of individual learner performance, the content and planning of revision lessons in this study were contingent on assessments of learner drafts, with the researcher drawing teaching points based on the more prevalent problems learners had demonstrated. It would hopefully be of interest to researchers and practitioners to see how learners reacted to this kind of feedback-embedded instruction and whether it enabled learners to improve their drafts. Before details of revision lessons are presented, the next section discusses recent literature on assessment and feedback and how they may facilitate teaching and learning.

LITERATURE REVIEW

Much of the recent literature on assessment, especially regarding formative assessment, emphasizes the educational aspect of assessment. According to Black and Wiliam (2009), assessment is beneficial to learning when it informs both learners and teachers of where the learners are going, where they are right now, and how to get the learners from where they are to where they want to go. Assessment of this nature is often called assessment *for* learning, so as to be distinguished from the more traditional concept of assessment *of* learning, in which the priority is assessment itself. A few suggestions on feedback practice that are conducive to assessment *for* learning are discussed below.

Full Communication and Opportunities for Dialogue

Research has shown that even the most carefully worded feedback often holds little meaning for students (Maclellan, 2001). Students either do not pay attention to the feedback or cannot understand it; even when they do, they often do not want to act upon it (Brookhart, 2007-2008). According to scholars, the question does not lie in the content of the message itself, but rather in whether the communication is successful. Price et al. (2010) illustrated how communication breakdown is prevalent when instructors send very concise and often obscure notations to students, limiting their professional opinions to the margins of the page, and especially when learners are not given a chance to ask for clarification. Therefore, opportunities for dialogue are critical if learners are to understand the feedback. Effective feedback has to be

delivered face-to-face with ample class time for learner questions and discussions, in order to ensure that the message is fully communicated.

Considering Learner Capacity and Adopting a Learner's-Eye View

Several studies (Hyland, 2000; Lee, 2009, 2011a; Williams, 2004) have suggested that much feedback is teacher-centered, leading to passive and dependent students. It is noted that corrective feedback is often solely focused on informing students of their errors. When learners receive their writing "awash in red ink" (Lee, 2007, p. 193), there can even be damaging psychological impacts. Student interviews in Lee's study have shown that they want to learn more about the criteria of good writing and are interested in trying other feedback options such as in-class discussions and conferences with teachers.

Echoing the notion of the possible damaging psychological impact of feedback (Lee, 2007), Brookhart (2007-2008) reminds teachers to address both cognitive and motivational factors in formative feedback. Moreover, Brookhart points out that a student can only hear the message when he is listening, when he can understand, and when he feels that it is useful to listen. In addition, in terms of the amount of feedback, research results indicate "less is more." But teachers often give too much and thus overwhelm students. It is also said that teachers should not only limit the amount of feedback but also prioritize areas of improvement for learners. This, coupled with the notion mentioned in the previous section of the effectiveness of the communication, lends itself to feedback that is limited and selective in content as well as thorough and illuminating in delivery.

A Balance Between Local and Global Issues

One branch of feedback research has investigated what types of information are included in typical instructor feedback. Lee (2007) investigated the nature of teacher feedback in Hong Kong writing classrooms. Based on a sample of 174 student texts and the accompanying 5,335 teacher feedback points, the author concluded that 94.1% of the points focused on form, 3.8% on content, and only 0.4% on organization. Similarly, in an American university, it was discovered that teacher feedback was oriented more towards local than global issues despite the fact that teachers reported and perceived that they were doing the contrary (Montgomery & Baker, 2007). Such lopsided emphases on local issues in teacher feedback has been the norm until very recently (Ferris, Brown, Liu, & Stine, 2011), despite continuous advice from composition researchers for practitioners to attend to a wide range of textual issues that include content and organization.

Other Feedback Principles

To help ensure students attend to and use feedback, Price et al. (2010) suggested that, instead of leaving learners to deal with feedback on their own and to wait for the somewhat distant next assignment in which they can apply the comments, opportunities for immediate use should be built into the design of tasks.

Nicol and Macfarlane-Dick (2006) linked assessment with self-regulated learning and pointed out that students have the ability to appraise feedback given to them. While learners' assessment ability often requires cultivation and practice, instructors can help learners develop this ability, and thereby transform learners from passive recipients of feedback to proactive users who are able to assess and lead their own learning. In this respect, the seven feedback principles conducive to learner self-regulation which Nicol and Macfarlane-Dick (2006) list are as follows: (1) help clarify what good performance is (goals, criteria, expected standards); (2) facilitate the development of self-assessment (reflection) in learning; (3) deliver high quality information to students about their learning; (4) encourage teacher and peer dialogue around learning; (5) encourage positive motivational beliefs and self-esteem; (6) provide opportunities to close the gap between current and desired performance; and (7) provide information to teachers that can be used to help shape their teaching. Numbers three through six coincide with the principles mentioned in previous sections in that effective communication and learner motivation for improvement are the keys.

THE STUDY

Purpose of the Study

The purpose of this study was to examine the effects of integrating feedback into revision lessons in the absence of providing individual feedback. It should be noted that, in this design, feedback is no longer treated as the final stage of homework assignments. Instead, it becomes a major part of the instruction between the drafting and revising stages. The instructor tried to direct her own attention towards diagnosing pervasive problems in students' writing and to design instructions that could assist learners to understand the problem and appropriately revise their work. But would it work? In the absence of specific individual feedback, would learners feel empowered or less secure? It was unclear if individual students would be able, after feedback-embedded revision lessons, to read their drafts critically, find similar problems in their own

work, and eventually improve their drafts. With these issues in mind, the researcher tested this approach with a group of college EFL learners in a regular course setting and examined its effectiveness.

The Context and Participants

The participants were 38 non-English-major college freshmen at a university in northern Taiwan. In Taiwan's junior and senior high schools, educational objectives are, in general, closely tied to the university entrance examination. For this examination, which almost all students must take, there are two writing components in the English section, in addition to multiple-choice questions. One is the translation of two compound or complex sentences from Chinese to English, comprising eight out of 100 total points for the English section. The other is a composition of one or two paragraphs totaling approximately 120 words. This piece is mostly narrative, such as describing the events depicted in a four-frame comic, for a total of 20 points. The remaining 72 points are allotted to multiple-choice questions on vocabulary, grammar, and reading comprehension.

Integrating Feedback into Revision Instructions

In this study, all learners had to complete three sets of writing tasks, each of which consisted of drafting and the follow-up revision. The author reviewed learner drafts and then taught revision strategies to learners, positioning the lesson between the draft and a mandatory revision. Since revision became a built-in component of the writing process, immediately following the revision instruction, learners had to revise their own work based on what they had learned. The rationale and instructional steps are summarized below. Figure 1 illustrates the procedures of this assessment-sensitive and feedback-embedded instructional plan, consisting of two phases: assessing the status quo in the first phase and moving to the goal in the second. The learners, with help from the teacher and their peers, were expected to subsequently take up the revision work on their own.

As shown in the figure, learners wrote a draft, which was followed by small group peer review. To model this review process, the teacher used a sample essay written by a more proficient peer (the teaching assistant for the course), and demonstrated how the quality could be evaluated using customized instructional rubrics. Students learned to assess and comment, and then applied what they had just learned in reviewing their peers' work. After class, the teacher collected learner drafts for assessment and prepared for the revision lesson to be given the following week. Rather than spending time circling all possible

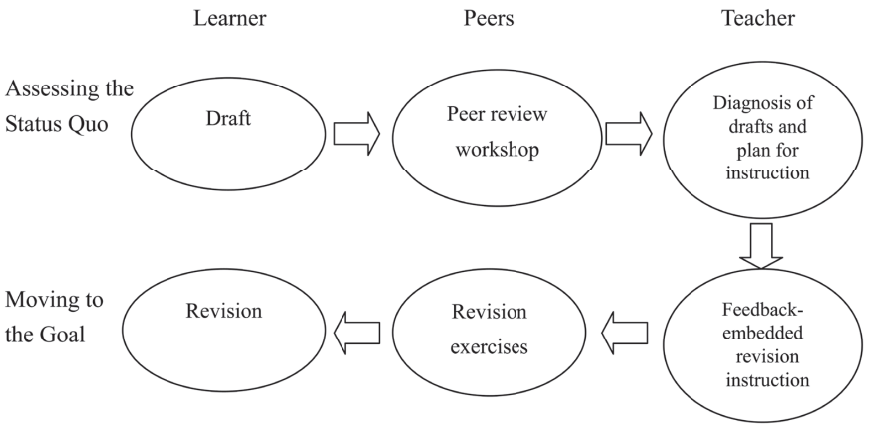

Figure 1
Instructional Procedure Flowchart

spotted errors and scribbling a great deal of condensed feedback in the margins, the teacher-researcher read drafts to diagnose where learners were in relation to the desired teaching outcome, and contemplated methods to assist learners to improve. She identified common strengths and problems, prioritized teaching points, chose representative student text passages, and designed specific revision exercises for lessons in the following week. The actual instructional points identified from learner drafts for three consecutive revision lessons are summarized in Table 1.

As illustrated in Table 1, teaching points progressed from the more general and basic aspects of revision and argumentation to some more specific areas. They reflected the common problems identified in learner drafts. For example, Point 3 in the lesson for Task 1 focused on supporting details and reminded learners to make sure that supporting details are directly relevant and support the main points. This point was made right after the basics in Points 1 and 2 because many details in learner drafts were not directly tied to and failed to support the main points. For more details of the revision instruction, a sample of the class handout is provided in Appendix A. It summarizes common arguments for both the pro and con sides from learner drafts and teaches a few strategies and major steps learners can follow to improve their drafts. Appendix B offers a learner draft excerpt chosen for class discussion and its subsequent revision.

In this design, feedback and instruction were no longer at two ends of a continuum, but rather intertwined in the middle (Hattie & Timperley, 2007). Feedback became the lesson and, unlike most

Table 1
Teaching Points in the Three Revision Lessons

Task 1 (Week 4 draft + Week 5 revision)	Task 2 (Week 7 draft + Week 8 revision)	Task 3 (Week 12 draft + Week 13 revision)
1. Explain initial procedures for revision, such as rereading own draft for an analysis of overall structure and argument points. 2. Check main points and examine the need for additions, deletions, or reorganization. 3. Check if supporting details support the main points. 4. Analyze sentence structure and check grammar.	1. Support with clear examples or explanations that are directly relevant. 2. Be careful when making assumptions about the background knowledge of readers; supplement with more details when necessary. 3. Check consistency of referents; be careful when switching among *I*, *we*, and *you*. 4. Discuss using dictionaries to help make decisions on word choice.	1. Make sure the discussion adheres to the topic; do not digress from the main topic. 2. Ensure good time management so there's enough time for a proper conclusion. 3. Do not bring up new points in the concluding paragraph. 4. Deletion can sometimes make the whole argument more focused. 5. Eliminate minor grammatical problems, such as subject-verb agreement, tense, and singular/plural forms.

traditional individual written feedback put at the margins, was communicated intentionally to the entire class face-to-face. This type of revision instruction, unlike Sengupta's (2000) and Butler and Britt's (2011) independent revision lessons, was contingent on assessment of and feedback to learner drafts. Based on more prevalent problems in the drafts, the instruction procedure was intended to help learners recognize necessary revision considerations and teach them concrete steps to improve an existing piece of work. It also communicated to learners the prioritized revision points, ways to improve their drafts, and the rationale behind the revising process. This instruction would hopefully have the potential to motivate learner revisions and improve writing quality. However, as could be reasonably imagined, it might fall short by not giving individual attention to learners and thus fail to improve learner drafts.

Procedures

The 38 participants were of various majors with varying levels of English ability. They were randomly assigned by school administration to one class for the required four-skill English course which met two hours a week for two eighteen-week semesters. By passing the two semesters, they are awarded four credits. The writing component of this particular course focused on opinion essays of about 300 words in length. Before the students were asked to do any writing, learning goals were communicated by discussing specific criteria and standards as well as viewing and evaluating multiple writing samples. These procedures, following the sequence depicted in Figure 1, were repeated three times in three writing tasks, each with a new topic. Using the blended design suggested by Ferris (2010), students wrote a draft, received a lesson on revision, then revised their draft. In weeks 4, 7, and 12, learners wrote their drafts on a given topic. In weeks 5, 8, and 13, learners were instructed on revision and revised their drafts. A brief course schedule, with the writing component in the above six weeks highlighted and other components in the remaining 12 weeks, is provided in Appendix C.

Research Questions

Under the research design presented above, the following specific research questions were posed:

(1) To what extent did student writing improve from drafts to the associated revised versions?

(2) To what extent did student drafts and the revised versions improve from one task to another?

(3) Based on self-reports, on what aspects did learners learn well, perform well, and have difficulties?

Data Collection and Analysis

In order to evaluate the writing quality of all drafts and revisions, the researcher invited two outside raters to grade the six pieces of learner work (three drafts and three revised versions). One of them had 18 years of experience as an EFL teacher as well as 10 years as a rater for a national standardized English proficiency test and the English composition section of the national college entrance examination. The other was a senior research assistant from the university's English department. Both were given the instructional rubrics that were used in class for discussion and peer review as a guide for their work. They were paid on a piece rate basis; however, neither was informed of

the purpose of this experiment during their work. Rating criteria, as illustrated in Appendix D, were adapted from the published criteria of the TOEIC writing component and the TOEFL iBT independent writing rubrics (Educational Testing Service, 2008a, 2008b, p. 14). Criteria included argument, organization, lexical use, grammar, and a composite holistic score. For instructional purposes, the original scale of 0 to 5 was expanded to a scale of 15.

Pearson correlation coefficients were computed to check the consistency between scores assigned by the two raters. Results (shown in Appendix E) indicated that both raters' scores on the three drafts and revisions in overall rating (H, holistic), argument (A), organization (O), lexical use (L), and grammar (G) were all positively correlated at a significant level. But inter-rater reliability was lower for the third article in terms of both argument and organization. After separate discussions with the two raters and re-reviewing the learner submissions, it became clear that the more experienced rater disapproved of many students' method of reasoning in the third task, referring to it as circular and offering support for the opposite side of the intended arguments. Therefore, she assigned these pieces much lower scores as compared to her previous standard and that of the other rater. Her reasons were somewhat subjective but valid. At this point, the researcher considered the possibility of introducing a third rater. After a discussion with statisticians at the statistics department on campus, it was decided that the author report the scores as they were, otherwise the validity of the scores might be contaminated through the introduction of a third rater. Therefore, caution was exercised when interpreting results related to argument and organization in the third article.

In addition to the above rating system, the cognitive experience associated with the feedback lessons as reported by learners was also a point of interest. The researcher wanted to know what students felt they had learned and not learned during this entire experience. In addition, the author was also curious about whether the lack of traditional written feedback for individual students caused any problems. A survey was therefore devised and conducted on the course Moodle platform at the end of the semester. Students responded anonymously in writing to a list of four open-ended short-answer items: (1) Please list three things you have learned about English writing and revision during the semester; (2) In terms of revising your own drafts, what was it that you did best; (3) In revising your own drafts, what was it that caused you the most difficulties; and (4) Any other relevant comments are welcome.

Each set of writing scores from the two raters on the six pieces of student work contained both a holistic score and four sub-scores.

First, inter-rater reliability was calculated. To answer the first research question, the two raters' averages were then used for paired-sample *t* tests between drafts and revisions. To answer the second research question, analyses of variance as well as multivariate analyses of variance were conducted for scores among drafts as well as among revisions. To answer the third research question, learners' short answers were categorized and tallied. Since no prescribed wording or choices were provided, students were free to use their own words to interpret and describe their individual learning experiences. Student responses formed a data set which was a collection of verbal descriptions listed as bullet points, sometimes coupled with explanations, and at other times containing multiple idea units. This had to be coded into categories and the results tallied. An analysis of the results is offered in the following section.

RESULTS

Descriptive statistics for the three writing tasks, together with results from paired-sample *t*-tests, are shown in Table 2. Mean scores improved from draft to revision for all three rounds, and the degree of improvement seemed to gradually level off from the first to the second and from the second to the third task. Comparing the sub-scores on each piece of work, it was found that learners generally scored higher on argument and lower on lexical use.

To answer the second research question regarding whether the drafts and revisions improved from one task to another, the author looked first at holistic scores, and further at the four sub-scores across the drafts and revisions. One-way analyses of variance were first conducted on holistic scores. For the drafts, the ANOVA was significant, $F(2, 90) = 5.134$, $p = .008$. Follow-up comparisons were conducted using Tukey's tests to evaluate pairwise differences among the means. There were significant differences in the means between drafts for the first and the third tasks, but not between drafts of the first and the second or the second and the third tasks. For the revisions, the ANOVA was insignificant, $F(2, 82) = 1.908$, $p = .155$. That is, the holistic scores suggest that the three revised versions did not improve over time. Further examination of the same question was made possible by comparing the four sub-scores among three drafts and three revisions. One-way multivariate analyses of variance were conducted. On drafts, the MANOVA was significant: Wilk's Lambda = .642, $F(8, 174) = 5.392$, $p = .000$. For the follow-up ANOVAs, the significance level was set at .0125 (.05 divided by 4, the total number of dependent variables). Difference

Table 2
Descriptive Statistics and *T*-Test Results on all Three Tasks

	Grading Criteria	Draft			Revision			Paired-Sample *t*		
		n	*M*	*SD*	*n*	*M*	*SD*	*t*	*df*	*p*
1st Writing	Holistic	**28**	**6.75**	**2.11**	**28**	**8.54**	**2.09**	**5.51**	**27**	**.000**
	Argument	28	6.81	2.10	28	8.61	2.02	5.84	27	.000
	Organization	28	6.12	2.27	28	8.27	2.01	5.95	27	.000
	Lexical Use	28	6.09	2.10	28	7.80	1.84	5.70	27	.000
	Grammar	28	6.10	2.25	28	7.75	2.03	5.43	27	.000
2nd Writing	Holistic	**31**	**7.63**	**2.22**	**31**	**8.98**	**2.17**	**4.16**	**30**	**.000**
	Argument	31	7.58	2.30	31	8.76	2.21	3.47	30	.002
	Organization	31	7.48	2.22	31	8.71	2.11	4.18	30	.000
	Lexical Use	31	7.11	2.12	31	8.34	2.12	3.67	30	.001
	Grammar	31	7.60	2.49	31	8.60	2.34	3.12	30	.004
3rd Writing	Holistic	**26**	**8.55**	**1.35**	**26**	**9.56**	**1.34**	**6.08**	**25**	**.000**
	Argument	26	8.28	1.51	26	9.44	1.40	5.85	25	.000
	Organization	26	8.26	1.42	26	9.31	1.23	4.84	25	.000
	Lexical Use	26	8.31	1.29	26	9.04	1.32	4.89	25	.000
	Grammar	26	8.56	1.48	26	9.12	1.42	3.49	25	.002

was found, again, between drafts of the first and the third tasks on all four sub-scores, with a 95% confidence interval of improvement on argument from 0.2 to 2.7, on organization from 0.9 to 3.4, on lexis from 1.1 to 3.4, and on grammar from 1.1 to 3.8. Improvements on the other pairs of the drafts were less consistent, with notable improvements from the first to the second restricted to organization and grammar, and from the second to the third restricted to lexis. For the revised versions, the MANOVA was insignificant: Wilk's Lambda = .851, $F(8, 156) = 1.660$, $p = .112$. As such, revised versions of the three tasks were not statistically different from one another. To summarize, although differences were spotted between drafts of the first and the third tasks, there was no reliable evidence for the author to claim improvement among drafts of the three tasks because of the low inter-rater reliability for part of the measures in the draft and the revised version of the third task.

To answer the third research question, the researcher first calculated the data size of the student survey responses (the students responded in Chinese, totaling 3,743 Chinese characters). A research assistant was later recruited to perform the same coding. Deviation occurred in 5.61% of the data set, which was resolved after discussion. Three hundred and twenty-one meaning units were identified, and 13 categories were devised. A summary of the most mentioned categories for the three questions is presented in Table 3 by number of counts. According to student reports, they learned more about textual features such as structure and organization, and argumentation and reasoning,

Table 3
Highlights of Student Reports on What They Had and Had Not Learned

Three Things Learned	One Thing Done Best	The Most Difficult Part
• Structure and organization (32) • Argumentation and reasoning (32) • Various meta-linguistic awareness (20) • Deletion (10) • Maintaining topic relevance (9)	• Structure and organization (21) • Argumentation and reasoning (18) • Correcting grammatical mistakes (11) • Refining word choice (7) • Length of texts (6) • Deletion (6)	• Limited vocabulary size (29) • Doubt on grammaticality (16) • Other uncertainties and confusions (8) • Argumentation and reasoning (8) • Length of texts (5) • Structure and organization (3) • Deletion (3) • Time pressure (3)

followed by various types of meta-linguistic awareness. Regarding what they felt they had done best at, students again predominantly mentioned structure and organization, followed by argumentation and reasoning, and then correcting mistakes and refining word choices. Their notes on difficulties centered mostly on lexis and grammar, followed by uncertainties regarding grammaticality and lexis. Argument and organization, while also mentioned, did not seem to stand out as a general problem for the majority of students. For the fourth question on additional comments, the most salient issues learners voiced were about not getting individual feedback from the teacher and not possessing self-confidence regarding what they chose to revise and whether the revisions could be considered successful.

DISCUSSION AND CONCLUSION

Before the discussion, two features in the design are worth noting. First, the experiment was not a one-shot study consisting of one draft and one revision only, but rather it followed Ferris' (2010) suggestion of a draft → feedback → revision → next draft sequence, which included both revision of the same text and the writing of new texts, over three rounds. Second, by including both holistic and four sub-scores assigned by two independent raters, the ratings helped us examine different aspects of learner writing.

The answer to the first research question was positive—all three revision scores improved as compared to their corresponding drafts on all five measures. Given the ambiguity and uncertainty involved for learners as they exercised discretion in applying what they had learned to their own drafts, this finding is significant. First, while it may not be a surprise for many to see that students' revisions improved on their drafts (it represented a second chance, after all), superior revisions are not a certainty. For example, Williams (2004) investigated revisions of students who had visited a writing center and found no measurable improvement. In her study, despite the one-on-one attention learners received, which is absent from most writing classrooms including the one depicted in this study, learners mainly transferred tutor suggestions verbatim as their revisions. Secondly, it was demonstrated in this study that, when teacher assessment and feedback was incorporated in the revision lessons and communicated to the student group face-to-face, students were able to revise their own drafts to an observable degree, even when individual feedback was absent.

The answer to the second research question, however, was not clear. Comparing the three drafts and three revisions, the only salient

difference was found between the drafts of the first and third writing tasks and not between other pairs. Although differences were observed in both the holistic and four sub-measures, the fact that one of the two raters disapproved of the argument rationale used by many students for the third task led her to assign much lower scores and consequently resulted in the low inter-rater reliability. This reality associated with human rating, though seldom reported in the published literature, makes the author more cautious about the results. More studies are certainly needed to clarify the problem.

For research question three, the data revealed that most students reported having learned and being confident about structure and organization. The aspects they claimed to do best at were generally consistent with what they reported having learned. Their difficulties were mainly associated with a limited vocabulary and uncertainty about grammaticality. The latter, as attested by some in the fourth open-ended question, pertained to the lack of individual feedback. These findings coincided with the pattern of sub-scores as shown in Table 2, where lexis and grammar scores were comparatively lower. On the other hand, argument, organization, and holistic scores were all higher. As for the degree of improvement from each draft to its revised version, students advanced the most in terms of organization. These results imply that, based on the revision lessons taught in this study, learners improved more at the global discourse level of writing and less on local lexical usage and grammaticality. Indeed, for these experienced college-level EFL students who were used to writing narratives, the organization of argumentative essays represented something new that could be picked up relatively successfully following a few rounds of instruction, whereas lexis and grammar issues require much more practice, as lexis and grammar are broad areas that need more than just a few rounds of instruction.

This study was different from Sengupta's (2000), in which individual feedback was added to revision lessons and this combination outperformed providing individual feedback alone. However, his revision lessons dealt with general principles of revising writing drafts, and the lessons were independent from his feedback prepared for individual learners. In this study, feedback was integrated into revision lessons in the hope that this face-to-face communication was more focused and full-fledged than written feedback and teacher time was prioritized for more common problems. Results showed that learners improved significantly from drafts to revisions, but there was no solid evidence to claim that the same improvement existed from one draft/revision cycle to the next. The learner survey showed that students

became more confident in the argument and organization of their writing, but were less so in lexis and grammar. Student concerns on the absence of individual feedback were also voiced.

In writing classes, where teachers are usually expected to give customized feedback to individual learners, the workload is multiplied by the number of students and further by the number of submissions. The number of students can often reach 30 or more in many classrooms such as those in Asia. Quite often these students have a homogeneous learning background and their learning needs are similar to a certain extent. Therefore, feedback points given may overlap to a great extent among different pieces of learner work. Once the more common points are chosen, organized, planned into lessons, and communicated face-to-face to allow for learner questions and teacher clarifications, these may address most learners' needs and are more easily communicated. This method can seem more challenging than teaching from a textbook because it involves spontaneity and contingency. However, it may not be so insurmountable since many major learner problems are similar and addressed in writing textbooks. What the teacher has to do is identify the needs through assessment and then teach to these needs.

The proposed approach in this study was a compromise under the constraint of resources and time. Revision instructions targeted at an entire student group, like most other classroom teaching, certainly do not cater to all learner needs. When learning histories and learner proficiency levels vary greatly, such lessons may not address the needs of the majority and run the risk of failure. Even in smaller classes where student learning needs are similar, teacher-student conferences allowing learners to discuss specific issues regarding his/her writing may be necessary to supplement this type of revision instruction.

In addition, the revision lessons were an initial attempt to incorporate feedback into this kind of instruction. They served more as general directions than concrete procedures to follow. Given the contingent nature of writing, there may never be concrete steps to follow, because contexts and learner populations vary. But once research of this type accumulates, it may be possible that certain principles and specific guidelines are recognized as central, which in turn will help this approach become more useful to teachers in the classroom.

LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

Partly because of its exploratory nature, this study has several limitations. First, there was no control group or other experimental

groups. More specifically, the participants were not compared with a group of learners who received individual feedback but not the revision lessons, or a group receiving both. What this study accomplished was to provide a scenario that was missing from Sengupta's (2000) study, i.e., the condition of teaching revision lessons which had integrated feedback for the particular batch of learner drafts. A more robust design is necessary if more substantive conclusions are to be made. Furthermore, caution should also be exercised in interpreting the results because of the lower inter-rater reliability on the third draft, which probably resulted from the absence of rater training.

Secondly, the kinds of revision on which students were rated were actually a combination of the revising and editing steps. In Butler and Britt's (2011), revising and editing are clearly distinguished. Experienced writers usually begin by revising globally for main ideas and structure, and then deal with local issues such as grammar and lexis. For this exploratory study on a new approach to providing feedback, such fine distinctions were not present. Future studies along this line may well take the different stages of revision and editing into consideration in the research design.

Another distinction not made concerns concepts and strategies associated with revision per se, as well as those of the argumentation genre. During the three feedback lessons, the instructor taught students some fundamental steps needed for revising drafts, and informed students of the structure and features of a good opinion essay. These two broad areas of instruction could be treated separately when learner needs differ. As mentioned earlier, Butler and Britt (2011) distinguished between a revision schema and an argumentation schema and prepared two different tutorials for each. Their findings indicated that each had significant impact on the quality of revision; but students who received both tutorials obtained no additive effect. While learners needed both kinds of instruction, it would be interesting to know if either could be more easily taught to students.

In conclusion, the current study proposed a method of integrating assessment of and feedback for learner drafts into revision instructions. The teacher diagnosed common problems and put feedback into revision lessons which focused on the why, what, and how of revision. Probably because of the better communication of messages in the customized revision lessons (rather than in written form), learners improved from drafts to revised versions. Although positive evidence was limited to the first research question, this rationale for providing feedback may warrant further investigation. In addition, learners reported having learned more about global features of organization and

argument as compared to local linguistic issues pertaining to lexis and grammar. This qualitative finding coincided with quantitative results and was consistent with previous studies. However, learners' expressed needs for individual attention reminded us to more carefully consider alternatives to the traditional type of feedback.

ACKNOWLEDGMENTS

The research reported in this paper was supported by Grant #100-2410-H-004-186-MY2 of Taiwan's Ministry of Science and Technology.

The author wishes to express her sincere gratitude to the editor, the copyeditor, and the anonymous reviewers for their professional assistance which helped improve the article from its earlier versions.

REFERENCES

Bailey, R., & Garner, M. (2010). Is the feedback in higher education assessment worth the paper it is written on? Teachers' reflections on their practices. *Teaching in Higher Education*, *15*, 187-198.

Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, *19*, 207-217.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*, 5-31.

Brookhart, S. M. (2007-2008). Feedback that fits. *Educational Leadership*, *65*(4), 54-59.

Butler, J. A., & Britt, M. A. (2011). Investigating instruction for improving revision of argumentative essays. *Written Communication*, *28*, 70-96.

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, *12*, 267-296.

Educational Testing Service. (2008a). *TOEFL® iBT test, independent writing rubrics (scoring standards)*. Retrieved from http://www.ets.org/Media/Tests/TOEFL/pdf/Independent_Writing_Rubrics_2008.pdf

Educational Testing Service. (2008b). *TOEIC speaking and writing sample tests*. Retrieved from http://www.ets.org/Media/Tests/TOEIC/pdf/TOEIC_sw_sample_tests.pdf

Ellis, R. (2010). Epilogue: A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition*, *32*, 335-349.

Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, *36*, 353-371.

Ferris, D. R. (2003). *Response to student writing: Implications for second language students*. Mahwah, NJ: Erlbaum.

Ferris, D. R. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing* (pp. 81-104). New York, NY: Cambridge University Press.

Ferris, D. R. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition*, *32*, 181-201.

Ferris, D., Brown, J., Liu, H., & Stine, M. E. A. (2011). Responding to L2 students in college writing classes: Teacher perspectives. *TESOL Quarterly*, *45*, 207-234.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*, 81-112.

Hyland, H. (2000). ESL writers and feedback: Giving more autonomy to students. *Language Teaching Research*, *4*, 33-54.

Lee, I. (2007). Feedback in Hong Kong secondary writing classrooms: Assessment for learning or assessment of learning? *Assessing Writing*, *12*, 180-198.

Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *ELT Journal*, *63*, 13-22.

Lee, I. (2011a). Feedback revolution: What gets in the way? *ELT Journal*, *65*, 1-12.

Lee, I. (2011b). Working smarter, not working harder: Revisiting teacher feedback in the L2 writing classroom. *The Canadian Modern Language Review*, *67*, 377-399.

Maclellan, E. (2001). Assessment for learning: The different perceptions of tutors and students. *Assessment and Evaluation in Higher Education*, *26*, 307-318.

Montgomery, J. L., & Baker, W. (2007). Teacher-written feedback: Student perceptions, teacher self-assessment, and actual teacher performance. *Journal of Second Language Writing*, *16*, 82-99.

Myhill, D., & Jones, S. (2007). More than just error correction:

Students' perspectives on their revision processes during writing. *Written Communication*, *24*, 323-343.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*, 199-218.

Polio, C., Fleck, C., & Leder, N. (1998). ''If I only had more time:'' ESL learners' changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing*, *7*, 43-68.

Price, M., Handley, K., Millar, J., & O'Donovan, B. (2010). Feedback: All that effort, but what is the effect? *Assessment & Evaluation in Higher Education*, *35*, 277-289.

Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, *35*, 535-550.

Sengupta, S. (2000). An investigation into the effects of revision strategy instruction on L2 secondary school learners. *System*, *28*, 97-113.

Silver, R., & Lee, S. (2007). What does it take to make a change? Teacher feedback and student revisions. *English Teaching: Practice and Critique*, *6*(1), 25-49.

Williams, J. (2004). Tutoring and revision: Second language writers in the writing center. *Journal of Second Language Writing*, *13*, 173-201.

Wingate, U. (2010). The impact of formative feedback on the development of academic writing. *Assessment & Evaluation in Higher Education*, *35*, 519-533.

ABOUT THE AUTHOR

Shu-Chen Huang has been teaching English at the college level and doing related research for nearly twenty years. Her research mainly focuses on English learning motivation and formative assessment. She has published in journals such as *System*, *Journal of English for Academic Purposes*, *International Review of Applied Linguistics*, *Language Testing*, and *Language Assessment Quarterly*.

APPENDIX A

Class Handout for Revision Instruction of Writing Task 2

I. Summary of arguments in learner drafts

Major Arguments for the Pro Side	Major Arguments for the Con Side
• Active participant > passive listener, knowledge internalized (learning theories) • College vs. high school—becoming more mature and independent • Interaction among learners, getting to know classmates • Becoming more expressive, part of learning • Group dynamics, leader/follower (organization behavior theories)	• Administrative procedures, waste of time, especially in large classes • Might be lack of real understanding for materials studies, teacher guidance needed • Wrong information/misperception spreading among learners • Free rider or discrepancy in contribution • Decreases teacher lecture time • Dominant/less dominant participants

II. Other possible argument points

1. Compromise
   - A healthy balance between the two
   - Teachers gradually prepare students for group discussion
2. You may also introduce some conditions by saying "It depends..."
   - On the nature of the subject area (math, business cases, etc.) being learned; On the age (type, proficiency level, etc.) of students; On the size of the class...; On student motivation (required vs. elective courses)
3. Sometimes forcing yourself to take only one side could be a great way to train critical thinking...

III. Instruction on major revision points

1. 解構自己的草稿，釐清各部分各自的功能與各部分相互間的關係，用文字清楚標示出來，如 My opinion is..., I have two reasons to support my position..., My first reason is..., Secondly, ..., Following that, I will provide one example..., To sum up..., etc.
2. 增加論述的內容、字數，在原來的論點上發揮、增加說明以支持論點。

3. 舉例要清楚，讀者可能沒有跟你一樣的經驗，所以來龍去脈要清楚交代 (when, where, who, what happened, why...)。
4. 檢視通篇中人稱的使用，修改不恰當、不一致的部分。
5. 思索原來想表達的意思，是否已充分表達出來？尋找恰當合適的字彙。

APPENDIX B

Samples of Learner Draft and Revision

Writing Prompt for Task 2

Many college teachers encourage group discussions among students. But some students feel listening to peers is a waste of time as compared to listening to teachers. Do you agree that teachers should encourage more group discussion? Include specific reasons or examples to support your answer.

An Excerpt of Learner Draft Sample

It is not unusual to know that there are people possessing the ideas that group discussion is only a waste of time. They may think that teachers' viewpoints towards the topic discussed is more accurate as well as professional than the shallow and inexperienced ones proposed by their peers. However, I argue otherwise. I am of the opinion that via the group discussion, we could have a different even brand new perspective toward the topic. Members in the group discussion are all special individuals, having their unique thoughts. Through group discussion, we could transform ourselves into a sponges, absorbing all the unprecedented ideas into our mind instead of being confined to the viewpoint offered by the teachers.

The Corresponding Revised Version

Many students consider group discussion a waste of time. They believe that teachers are an authoritative figure in the classroom, having more "accurate" and "professional" perspectives toward the subject matter being learned. In contrast, viewpoints proposed by their peers could hardly have much insight and are likely to be quite superficial. However, I argue otherwise. I am of the opinion that through group discussion, students could obtain different, or even brand new, perspectives for the same topic. Each member in the group is a unique individual and, with specific personal experiences, he or she may generate thoughts that simply have not come across other students' minds. Therefore, an open-minded learner does not have to be confined to the viewpoints offered by the teacher and has a great opportunity to absorb all the various ideas from his/her peers.

APPENDIX C

The Course Schedule

Week	Topics and Activities
1	Introduction to the material/instructor/course; Getting to know each other
2	Introducing e-platform; L2 writing experience reflection; Using course materials to your advantage
3	Reading aloud practice; Oral summary and discussion of the articles assigned; Introduction to some standardized test writing and instructional rubrics
4	Writing task 1; Modeling peer review with TA sample; Peer review workshop
5	Revision instruction 1; Revising 1st draft
6	1st exam on course materials; Practice and preparation for 1st oral exam
7	Writing task 2; Modeling peer review with TA sample; Peer review workshop
8	Revision instruction 2; Revising 2nd draft
9	Touring self-access learning center; Reading and discussing articles; Class visit by a foreign student from Nicaragua
10	1st oral exam in the lab
11	2nd exam on course materials; Practice answering one-minute questions
12	Writing task 3; Modeling peer review with TA sample; Peer review workshop
13	Revision instruction 3; Revising 3rd draft
14	Group oral reports; 3rd exam on course materials
15	Group oral reports; Wrap-up of the entire writing experience
16	Group oral reports; Reading and discussing articles
17	2nd oral exam in the lab
18	4th exam on course materials

APPENDIX D

Rubrics for Instruction and Grading

項目	Scoring Criterion/ Level	Excellent (90% & up) 15, 14, 13	Good (80%-89%) 12, 11, 10	Fair (70%-79%) 9, 8, 7	OK (60%-69%) 6, 5, 4	Poor (below 60%) 3, 2, 1
總分	Holistic Score 作文整體品質	極優秀	優良	良好	尚可	一般水準之下
Content 內容	Argument 主題／主張	+	主題清晰、立場明確、有解釋、很容易就知道作者的主張			
		(–)	找不到主題、或需費力尋找、令人不清楚作者的意圖			
	細部論點	+	有明確清楚的細部理由、解釋清楚、能說服人、適當舉例			
		(–)	沒有支持主張的理由、沒有充分說明理由、不能說服人			
	Organization 整體組織結構	+	有清楚開頭、逐漸發展完整的內容、良好收尾、整篇文章統一、連貫、順暢、有邏輯			
		(–)	漫無目標、雜亂的內容，或空有開頭結尾的形式			
Language 語言	Lexical Use 字彙／用語	+	用語恰當、自然、有變化、生動、有力			
		(–)	用語令人迷惑不解、重複多、不自然			
	Grammar 文法句構、書寫體例	+	句構、單複數、人稱、時態、動詞變化、標點符號、拼字等文法均正確、幾乎沒有錯			
		(–)	很多明顯的文法問題			

APPENDIX E

Pearson Correlation Coefficients of Scores between Two Raters

		Drafts					Revisions				
		H	A	O	L	G	H	A	O	L	G
1	*r*	0.79	0.71	0.59	0.75	0.75	0.88	0.94	0.81	0.65	0.72
	p	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
2	*r*	0.94	0.95	0.86	0.83	0.81	0.96	0.92	0.85	0.80	0.70
	p	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
3	*r*	0.56	0.33	0.57	0.48	0.37	0.40	0.23	0.35	0.63	0.50
	p	(0.00)	(0.08)	(0.00)	(0.01)	(0.05)	(0.04)	(0.26)	(0.08)	(0.00)	(0.01)

Note. 1 = 1st piece of writing, 2 = 2nd piece of writing, 3 = 3rd piece of writing. H = holistic, A = argument, O = organization, L = lexical use, G = grammar.

將評量回饋意見融入寫作修改教學以促進學生文稿之修改

摘 要

外語教師常對學生習作草稿給予個別書面回饋，這類回饋常未能使學生改好文章，部分原因在於給評的方式不能充分溝通訊息。近來有研究直接教學生修改，有效提升了學生文稿的品質。本研究融合上述兩種方法，將針對學生草稿的評語整合到寫作修改教學之中，整理出寫作修改教學內容，教學後由學生修改文稿，修改前後的文稿邀請兩位評分員獨立給分，另以開放式問題於課程平臺蒐集學生對此學習經驗的意見。研究結果顯示學生每次的修改都較之前的草稿有顯著的進步，其中文章內容及架構的進步幅度優於文法及用字，然而沒有資料可證明學生在進入下一回的寫作及修改練習時，亦有同樣明確的進展。

關鍵詞：書面回饋　寫作修改教學　促進學習的評量