行政院國家科學委員會專題研究計畫 期末報告

學習動機、課室評量與教師權變(第 2 年)

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 100-2410-H-004-186-MY2
執 行 期 間 ：101 年 08 月 01 日至 102 年 07 月 31 日
執 行 單 位 ：國立政治大學外文中心

計 畫 主 持 人：黃淑真
共 同 主 持 人：崔正芳
計畫參與人員：助教-兼任助理人員：劉珮宜
助教-兼任助理人員：李承泰
博士班研究生-兼任助理人員：徐維里

報 告 附 件 ：移地研究心得報告
出席國際會議研究心得報告及發表論文

公 開 資 訊 ：本計畫涉及專利或其他智慧財產權，2 年後可公開查詢

中 華 民 國 102 年 10 月 04 日

中 文 摘 要 ： 本計畫深入研究學習動機與課室評量的關聯，及將課室評量與教師權變結合的方法。計畫植基於研究者 96 年度國科會計畫成果之兩篇論文，其一探討師生對以考試與分數激勵學習的看法、其二比較考試與報告二種不同評量方式，對學生動機及學習策略的影響。二文結論指出，學生動機會影響其因應課堂要求的態度，而課室評量的安排亦影響學習動機；教師在安排課程配分給分時，多會將學生動機與其可能投入的程度納入考量。本計畫前半部將擴充對師生做法、態度與信念的調查，以檢驗之前觀察到的本地考試文化效應以及師生心態，究竟為特屬於華人社會的現象，抑或是亞洲現象、跨國跨文化現象？藉由之前的研究基礎，擴大問卷調查與訪談的規模，梳理出動機與評量之間微妙的關係、其背後的成因，以及進一步提出動機與評量理論修正的方向，和可供大專英語教師參考的做法。計畫的後半部以最新的課室評量理論（Black & Wiliam, 2009）為本，據此設計符合近年 assessment FOR learning「評量以助學習」思潮的大專生英語說寫教學流程。同時依新理論所提的 teacher contingency 觀念，帶進企業管理的權變理論，並檢驗教師權變的可行性與效益，將以準實驗研究法設計說寫教學各兩套流程，控制組依一般預擬教案在學生練習後介入，實驗組之不同處在教案不預設，而將於教師閱覽過學生作品後才決定。實驗前後蒐集學生動機問卷，並進行作品修改前後的內容分析，希望結果有助於探討教師權變的可能與修正方向。動機與評量、權變的結合是英語教學的創新思維，其理論與實用價值均值得投注
更多心力。

中文關鍵詞： 形成性評量、課室評量、學習動機、教師權變、英文寫作、回饋、修改

英 文 摘 要 ： Traditional teacher feedback given in response to L2 students' writing has been characterized as laborious, ineffective, and even harmful. Recent studies have recommended various promising methods that help to remedy observed problems. In a required integrated EFL course for a group of 38 college freshmen, the researcher implemented several suggested feedback practices to replace conventional written feedback on individual submissions. Major non-traditional variations included revision lessons between the drafting and revision stages, with teaching points based on pervasive patterns

identified from learners' drafts. After three rounds of writing, with each round consisting of a draft-instruction-revision sequence, the three drafts and revisions were evaluated using five items scored by two independent raters. Also, learner experience was clarified based on responses to an open-ended questionnaire. Findings indicate that learners improved from drafts to revisions for all three tasks on all five measures. Among the three drafts and three revisions, significant differences were found between the first and the third drafts. Furthermore, learners reported that they improved more on organization and argument than on local linguistic features. In sum, the proposed assessment-based feedback lesson demonstrates its potential value as a replacement for written feedback provided on individual student papers.

英文關鍵詞： formative assessment, teacher contigency, written feedback, corrective feedback, second/foreign language writing, revision, assessment for learning

EFL Writing Instruction Based on Classroom Assessment to Motivate Learner Revision

Table of Contents

List of Tables

List of Figures

Introduction

Revision is part of the writing process for all writers; students engaged in second language writing especially need to cultivate the ability to revise and improve their own work. Numerous studies have discussed various approaches where teachers help learners revise, such as error feedback (Ferris, 2010), trained peer review (Min, 2006), explicit instruction of revision strategies (Sengupta, 2000), teacher-student conferences (Ewert, 2009), and tutoring in a writing center (Williams, 2004). That said, written feedback is probably the most commonly adopted approach—as such, it has received the majority of research attention. Moreover, there is evidence suggesting a significant positive relationship between the teachers' ability to provide quality feedback and gains in their students' writing achievement (Parr & Timperley, 2010). Yet interestingly, research has also 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, 2007/8).

One branch of feedback research investigates 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 actual student texts and the accompanying 5335 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 and Stine, 2011), despite continuous advice from composition researchers for practitioners to attend to a wide range of textual issues

that include content and organization.

In addition to the feedback content, many researchers have focused on the effectiveness of teacher feedback in improving student writing. A number of studies have provided evidence for the efficacy of written corrective feedback, especially in terms of the accuracy of certain well-defined linguistic features (e.g. Ferris, 2006; Bitchener & Knoch, 2010), although skepticism still persists (Truscott, 1996; Truscott & Hsu, 2008). For example, it was found that direct correction was fastest in producing accurate revisions; however, despite the significant quick effect, actual learning was debatable in comparison with the more time-consuming learner self-correction process (Chandler, 2003). Moreover, there are doubts about the ecological validity of feedback studies focusing exclusively on grammatical accuracy (Ellis, Sheen, Murakami, & Takashima, 2008) in terms of whether studies of this kind are pedagogically relevant to or credible for writing teachers (Bruton, 2009).

Some teachers believe that feedback on students' writing should focus on more than linguistic accuracy and include suggestions regarding content and organization, or even take into consideration students' cognitive and affective development. However, empirical findings for feedback effects on these other non-linguistic aspects tend to be somewhat negative. For example, several studies (Hyland, 2000; Lee 2009; Lee, 2011; Williams, 2004) have suggested that most feedback is teacher-centered, leading to passive and dependent students. Another report of negative effect on affect noted that corrective feedback is often solely focused on informing students of their errors. When learners receive their writing "awash in red ink", there can even be damaging psychological impacts (Lee, 2007). Student interviews in the study have shown that students 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.

Further, issues associated with providing feedback on student submissions are not restricted to writing teachers only. In many subject disciplines at different levels of education, feedback is an indispensible, but mostly peripheral, part of teacher practice. Recently, several studies (Bailey & Garner, 2010; Price, Handley, Millar, & O'Donovan, 2010) have questioned the common practice of written feedback, especially in view of the amount of time and effort required by teachers. Bailey and Garner (2010) interviewed 48 lecturers in the British higher education context and found that these lecturers were confused about what feedback should achieve and what students do about it. The researchers referred to feedback as "having a Cinderella status on the margin of institutional structures and processes" (p. 187). The majority of the "best feedback approach" suggestions in the second language writing literature also demand a great deal of teacher time and effort (Ferris, 2003). In addition, some researchers have attested that the undergraduate students pay little attention to teacher feedbacks or do not understand nor act upon them (Sadler, 2010; Wingate, 2010), although the small number of those who did pay attention and act on the suggestions improved in the areas previously criticized. Two reasons for students' non-engagement with feedback included failure to offer any strategies for using feedback (Silver & Lee, 2007), and students' low motivation or low self-perception as writers, factors which are usually neglected in second language acquisition studies (Ellis, 2010). These authors also suggested that teachers and researchers pay close attention to the ways feedback content is communicated as well as learners' affective factors, which will increase the likelihood that students utilize feedback.

Several studies have investigated alternatives to abridged written corrective feedback when teaching revision and obtained positive results. First, 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 had a significant effect on writing scores, which suggests that such explicit instruction may facilitate the development of an awareness of discourse-related features in L2 writing. Another study with similar objectives was situated in first language tertiary context in the US: Butler and Britt (2011) found their students underprepared for academic writing, such that they could not write well-structured arguments. In response, they designed two writing tutorials to help students revise their argumentative essays—an argument tutorial and a global revision tutorial. Students could complete the tutorials independently, and both were shown to help improve the revised submissions; interestingly, the improvement associated with completing both tutorials did not exceed that associated with completing either of the individual tutorials. In contrast, learners who did not complete either of the tutorials made more local changes, and their revisions were generally not considered to be much improved as compared to their drafts. The authors also discussed the potential pitfalls of building an enduring misconception in students if teachers continue to provide feedback on local errors. In sum, attempts to direct learner attention beyond the local issues of revising a draft have been encouraging.

Recent literature on learning assessment has also provided insights on making feedback more effective (e.g. Black & Wiliam, 2009). First of all, 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. Second, Price et al. also 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, especially when learners are not given a chance to ask for clarification. Opportunities for dialogue are critical if the feedback is to be understood by learners. Thirdly, echoing the notion of the possible damaging psychological impacts of feedback (Lee, 2007), Brookhart (2007/8) reminded teachers to address both cognitive and motivational factors in formative feedback. Moreover, Brookhart pointed 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 amount of feedback, research results indicated “less is more.” But teachers often give too much and overwhelm students. It is also said that teachers should not only limit the amount of feedback, but also need to prioritize areas of improvement for learners. Finally, instead of looking at surface errors and mistakes of second language writers, McGarrell and Verbeem (2007) advocated that teachers use “an inquiring stance” in constructing feedback: including informational questions on early drafts may help learners clarify their communicative intentions and negotiate emerging meanings, thereby guiding learners as they revise and refine their drafts.

A similar set of suggestions were recommended by Nicol and Macfarlane-Dick (2006), who pointed out that students have the ability to appraise feedback given to them. They also linked assessment with self-regulated learning. While learners’ assessment ability often requires significant improvement, 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, seven feedback principles conducive to learner self-regulation 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 teaching. Numbers three through six coincide with the suggestions mentioned above, in that effective communication and learner motivation for improvement are the focal point.

Replacing Individual Written Feedback with a Face-to-face Feedback Instruction to All

It is difficult to imagine applying the above feedback guidelines when feedback is delivered individually to students in the written form as a supplement to homework grades, as is traditionally practiced and reported in the literature. In this study, the author made feedback the center of L2 writing instruction, positioned between the first draft and a mandatory revision. The detailed rationale and instructional steps (Huang, under second review) are beyond the scope of the current study—they are briefly summarized below.

As shown in Figure 1, learners wrote an initial draft, which was followed by group peer reviews. To model this review process, the teacher would use a sample written by a more proficient peer (the teaching assistant for the course), and demonstrate how the quality could be evaluated using customized instructional rubrics. Students subsequently learned how to assess submissions in terms of areas for improvement and then reviewed their peers' work.

The more intensive teacher evaluations began once learner drafts were collected. Rather than spending time circling all possible spotted errors and scribbling a great deal of condensed feedback in the margins, the teacher researcher read each draft to diagnose where learners were in relation to the desired teaching outcome, and contemplated effective methods to assist learners to improve. She identified common

strengths and problems, prioritized teaching points, chose representative student text chunks, and designed specific revision exercises as the basis for follow-up revision lessons. As such, feedback and instruction were no longer at two ends of a continuum, but rather intertwined in the middle (Hattie & Timperley, 2007). During the following class meeting, a significant amount of time was allotted to the selected feedback points, and two-way dialogue between the teacher and students was encouraged to check understanding.

It is believed that such feedback instruction has much greater potential to motivate learner revisions. These types of assessment-based instruction and exercises help learners recognize necessary revision considerations and concrete steps to improve an existing piece of work. More importantly, revision becomes a built-in component of the writing process, as immediately following the feedback instruction learners are invited to revise their own work based on what they just learned.

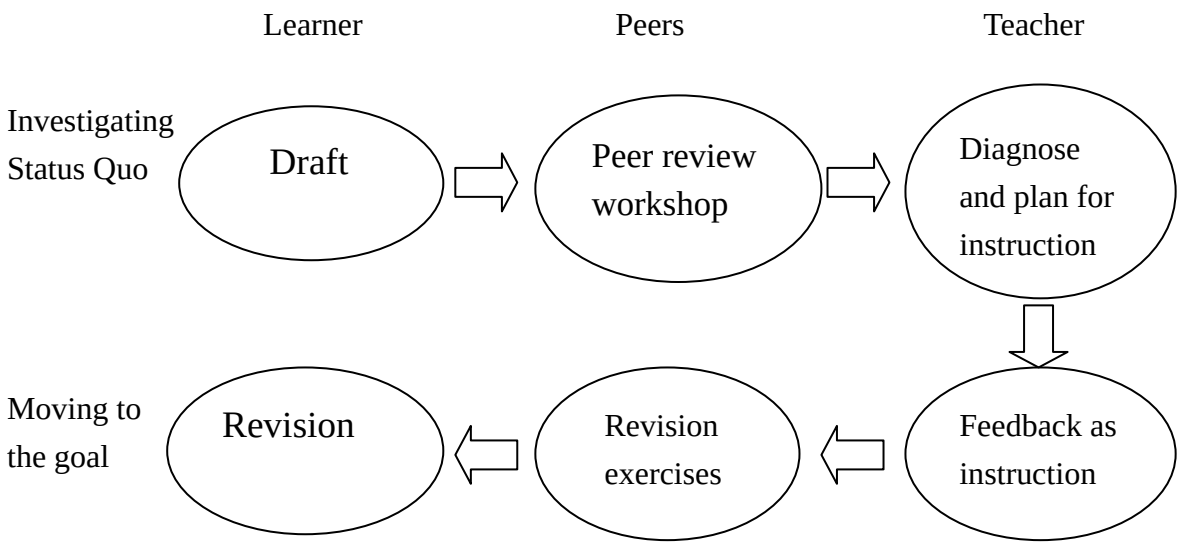


Figure 1. Sequence of an assessment-based L2 writing instruction process

The main thrust of this design lay in replacing the traditional written feedback given to individual assignments with a more purposive revision lesson based on actual

student drafts, aimed at clearly communicating to learners the prioritized revision points, ways to improve their drafts, and the thought processes behind the revision.

The Study

The purpose of this study was to examine the effects of the above feedback process, which incorporated recently developed feedback principles into a revision lesson delivered to an entire student group. It should be noted that following these principles ensures that feedback is no longer treated as the final stage of homework assignments—not to be discussed after it is provided. Instead, it becomes a major part of the instruction between the drafting and revising stages. Instructors no longer are required to laboriously scribble comments on individual submissions—a practice that has been criticized as ineffective and even harmful. Instead, attention is strategically directed towards diagnosing pervasive problems in learner work and designing instructional tools that can assist learners to appropriately revise their work. This type of approach goes beyond the traditional steps associated with writing instruction pedagogy. But does it work? In the absence of specific individual feedback, it is unclear whether learners would feel empowered or less secure than before. Moreover, questions persist regarding whether students are able to read their drafts as critically as a teacher does, find problems similar to those that are traditionally shown to them, and eventually make useful revisions. With these issues in mind, this preliminary investigation was initiated by administering the proposed teaching approach to a group of college EFL learners in a regular course setting and collecting data to examine the effectiveness.

The Context and Participants

The participants consisted of a group of college freshmen in Taiwan. Prior to entering college, students in Taiwan have learned English as a foreign language since the third grade in primary school. Many start even earlier or get extra hours of exposure outside of school. Primary and secondary school education is conducted under guidelines and a selection of textbooks approved by the Ministry of Education. In junior and senior high, educational objectives are in general closely tied to the entrance examination into the subsequent level. For the joint entrance examinations into college, which almost all students must take, there are two written parts in the English section: one is the translation of two compound or complex sentences from Chinese to English (eight out of one hundred total points for the English section) and the other is a composition of one or two paragraphs totaling approximately 120 words; this piece is predominantly narrative, such as describing a series of four comics (twenty points).

Procedures

The 38 participating freshmen were of various majors enrolled in a four-skill required English course, which met two hours a week for two eighteen-week semesters; after passing the course they were awarded four credits. The writing component of this course focused on opinion essays of about 300 words in length. Before being asked to do any writing, learning goals were communicated by discussing specific criteria and standards as well as viewing and evaluating multiple writing samples. The writing lesson, following the sequence depicted in Figure 1, was repeated three times—each with a new topic. Using the blended design suggested by Ferris (2010), students wrote a draft, a revision (after the teacher's feedback lesson, where the feedback was not written or exclusively corrective), and then a new text. In

weeks 4, 7, and 12, learners wrote three drafts based on new topics. Each 30-minute timed draft was followed by the teacher-led discussion of a TA sample and learners' small group peer reviews. Drafts and peer review results were studied by the instructor between the two weekly meetings of one writing unit. In weeks 5, 8, and 13, feedback was given to the entire class as a revision lesson for 50 minutes before learners started to revise for another 30 minutes in the second period. It was in this first 50-minute revision lesson where the collective wisdom of the feedback literature was tested. The teacher briefed diagnostic summaries, identified selected problems, demonstrated revisions, and eventually presented problems using learner excerpts, which students initially solved in small groups; later, the teacher led a class discussion to assess the proposed solutions. In the second 50-minute session, learners were given 20 minutes to reread their own first drafts, review peer evaluation results from the previous week, self-assess using the same rubrics, and set a revision goal based on a revision checklist provided by the teacher. The remaining 30 minutes were reserved for revising the draft. The three writing topics and the content of the revision lessons are depicted in Appendix A. Throughout the semester, each student was required to produce a total of six pieces of writing: three drafts and three revisions.

Research Questions

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

1. Did student writing improve from drafts to revisions?
2. Did student drafts and revisions improve from one task to another?
3. What did learners report regarding their learning?

Data Collection and Analysis

In order to evaluate the writing quality of all drafts and revisions, two outside raters were separately invited to grade the six pieces of learner work. 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 part 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 (see Appendix B) 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 B, was adapted from the publicized criteria of the TOEIC (Test of English for International Communication) writing component. Components included argument, organization, lexical use, grammar, and a holistic score, each on a scale of 1 to 15.

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. The data set was a collection of verbal descriptions listed as bullet points, sometimes coupled with explanations, and other times containing multiple idea units. A detailed analysis of the results is offered in the following section.

Results

Pearson correlation coefficients were computed to check the consistency between scores assigned by the two raters. Results in Table 1 indicate 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, although inter-rater reliability was lower for the third article in terms of both argument and organization. After separate discussions with two raters and re-reviewing the learner submissions, it became clear that one rater disapproved of many students' method of reasoning in the third submission, referring to it as circular and offering support for the other side; 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 certainly valid. At this point, the researcher

considered the possibility of introducing a third rater. After a discussion with statisticians, it was decided to retain the original scores, since the correlations were all positive and the majority of figures were at a satisfactory level. Moreover, the validity of the scores may have been contaminated through the introduction of a third rater at this point. That said, caution was exercised when interpreting results related to the argument and organization in the third article.

Table 1. Pearson correlation coefficients of scores between two raters

		Drafts					Revisions				
		H	A	O	L	G	H	A	O	L	G
1	r	0.789	0.712	0.590	0.750	0.749	0.882	0.939	0.807	0.645	0.721
	p	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
2	r	0.943	0.947	0.863	0.828	0.805	0.963	0.923	0.848	0.803	0.699
	p	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
3	r	0.563	0.330	0.567	0.481	0.367	0.397	0.229	0.348	0.631	0.496
	p	(0.001)	(0.080)	(0.001)	(0.008)	(0.050)	(0.044)	(0.260)	(0.082)	(0.001)	(0.010)

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

Descriptive statistics for the three drafts and revisions, together with results from paired-sample *t*-tests, are shown in Table 2; they are also depicted in a line graph in Figure 2. It can be seen from Figure 2 that the mean scores improved from the drafts to revisions 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. The second and third drafts received lower scores than their previous revisions, but they both received higher scores than the previous drafts. Comparing the five scores on each piece of work, it was found that holistic and argument scores were generally higher while grammar and lexical use scores were lower.

Table 2. Descriptive statistics and *t*-test results on all three writings

	Grading	draft			Revision			paired-sample *t*		
	Criteria	*n*	*M*	*SD*	*n*	*M*	*SD*	*t*	*df*	*p*
1st	**Holistic**	**32**	**6.73**	**2.29**	**28**	**8.54**	**2.09**	**5.509**	**27**	**.000**
Writing	**A**rgument	32	6.75	2.21	28	8.61	2.02	5.837	27	.000
	Organization	32	6.08	2.29	28	8.27	2.01	5.954	27	.000
	Lexical Use	32	6.06	2.12	28	7.80	1.84	5.696	27	.000
	Grammar	32	6.03	2.39	28	7.75	2.03	5.426	27	.000
2nd	**Holistic**	**32**	**7.61**	**2.30**	**31**	**8.98**	**2.17**	**4.160**	**30**	**.000**
Writing	**A**rgument	32	7.61	2.32	31	8.76	2.21	3.473	30	.002
	Organization	32	7.47	2.25	31	8.71	2.11	4.184	30	.000
	Lexical Use	32	7.09	2.11	31	8.34	2.12	3.671	30	.001
	Grammar	32	7.59	2.52	31	8.60	2.34	3.115	30	.004
3rd	**Holistic**	**29**	**8.43**	**1.43**	**26**	**9.56**	**1.34**	**6.080**	**25**	**.000**
Writing	**A**rgument	29	8.22	1.52	26	9.44	1.40	5.847	25	.000
	Organization	29	8.24	1.43	26	9.31	1.23	4.838	25	.000
	Lexical Use	29	8.29	1.31	26	9.04	1.32	4.888	25	.000
	Grammar	29	8.48	1.54	26	9.12	1.42	3.486	25	.002

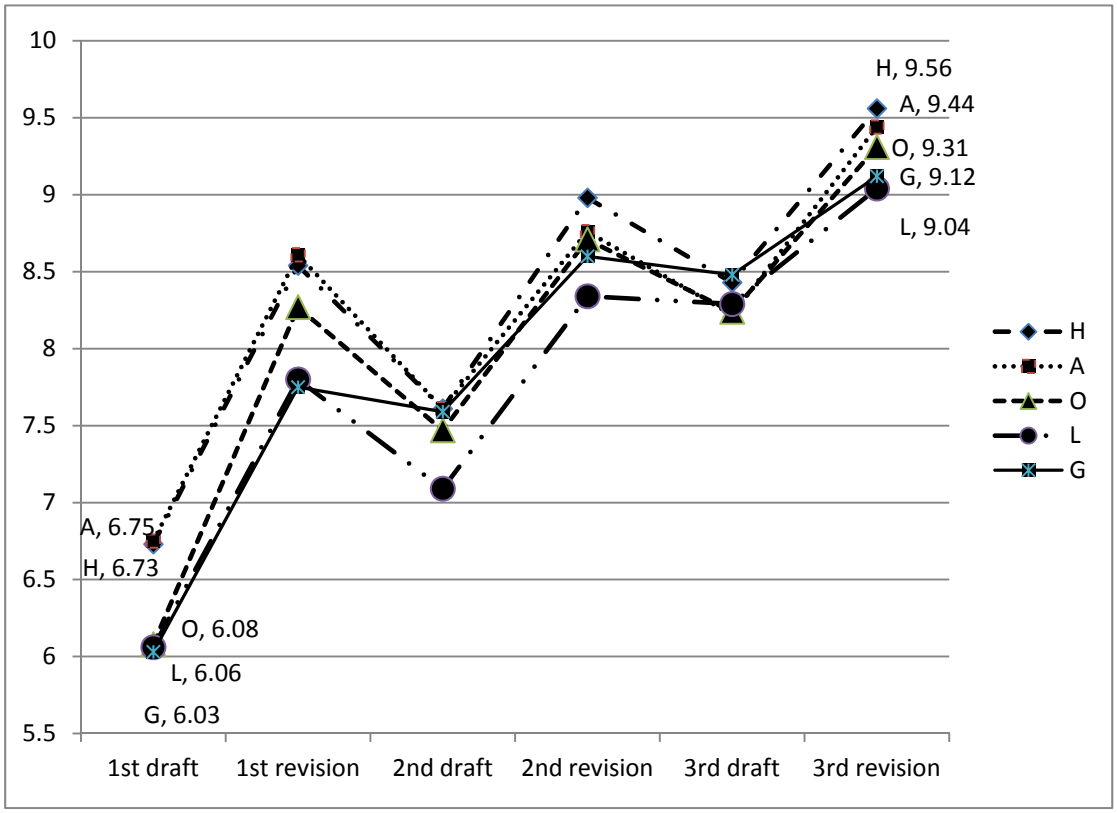


Figure 2. Average holistic and sub-scores in all six pieces of writing

Note: H – holistic; A – argument; O – organization; L – lexical use; G – grammar

To answer the first research question, paired-sample *t*-tests were performed between drafts and revisions for all three tasks. In order to control for Type I errors, the significance level was adjusted and set at .05 divided by 3. The results, as shown in the rightmost columns of Table 2, indicate that all revisions were significantly better than their drafts in terms of both holistic scores and all four sub-scores. The answer to the first research question was therefore positive.

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 firstly 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 the first and the third drafts, but not between the first and the second or the second and the third drafts. 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 was found, again, to be between the first and the third draft on all four sub-scores, with a 95% confidence interval of improvement on argument from 0.2 and 2.7, on organization from 0.9 and 3.4, on lexis from 1.1 and 3.4, and on grammar from 1.1 and 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 revisions, the MANOVA was insignificant: Wilk's Lambda = .851, F (8, 156) = 1.660, p = .112. As such, the three revisions were not statistically different from one another.

To answer the third research question, the researcher first calculated the data size (totaling 3743 Chinese characters). Among them, 321 meaning units were identified and 13 categories were induced. A summary of the most mentioned categories for the three questions are listed in Table 3 by number of counts. Based on student reports, they learned more about textual features such as structure and organization, and argumentation and reasoning, followed by various types of meta-linguistic awareness. Regarding what they had done best at, students again predominantly mentioned structure and organization, followed by 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.

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	- Structure and organization (21) - Argumentation and reasoning (18) - Correcting grammatical mistakes (11) - Refining word choice	- Limited vocabulary size (29) - Doubt on grammaticality (16) - Other uncertainties and confusions (8) - Argumentation and reasoning (8)

relevance (9)	(7) - Length of texts (6) - Deletion (6)	- Length of texts (5) - Structure and organization (3) - Deletion (3) - Time pressure (3)

Discussion

Before the discussion, I would like to highlight the uniqueness of this study. Unlike most written corrective feedback studies, the feedback here was neither written nor corrective. Instead of giving individual feedbacks, the teacher's effort was directed to preparing an in-depth feedback for all, which became the major body of instruction for revision. This design was mainly a product of incorporating recent findings from feedback studies in L2 writing and from feedback studies in the area of formative assessment (or assessment for learning). While many L2 writing teachers consider not giving individual feedback irresponsible or unethical (Bruton, 2009), this study experimented with abandoning individual feedback altogether, rather than loading the teacher with providing both individual feedback and an additional feedback lesson for all students, as was done in the Sengupta (2000) study. Moreover, two features in the design are worth noting. First, the experiment was not a one-shot consisting of one draft and one revision only, but rather it followed Ferris' (2010) suggestion of a 'draft-feedback-revision-next text' sequence, which included both revision of the same text and the writing of new texts, over three rounds. This methodology made it possible to examine the effects of such feedback instruction not just on immediate revisions, but also on learners' uptake demonstrated in the subsequent new tasks. Second, the outcome measures adopted were similar to real world practices in terms of how learner writings are generally evaluated on standardized tests or in L2 writing classrooms. In the literature, some lab studies

focused very narrowly on specific grammatical features like the use of definite and indefinite articles and presented the change in number of such mistakes as an indicator of improvement (e.g. Bitchener & Knoch, 2010), with few implications for practicing teachers. Others used only impressionistic holistic ratings (Sengupta, 2000) or allowed discussions between raters, leading to concerns about the validity of the writing quality scores. In the current study, the ratings were made more comprehensive by including both holistic and four sub-scores generated by two independent raters, allowing for discrepancies to be recognized, which more closely resembles real-life practice. Hence, both ecological validity and pedagogical relevance (Bruton, 2009; Ellis et al., 2008) were considered in this study.

The answer to the first research question was an unconditional yes—all three revision scores improved as compared to their corresponding drafts for all five measures. Given the ambiguity and uncertainty involved in the learners' need to exercise discretion by individually applying the instructor feedback 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 were not a sure thing. 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, they mainly transferred tutor suggestions verbatim as their revisions. Secondly, while many teachers and students believe that individual feedback is a must, it was demonstrated in this study that students were still able to revise their own drafts to an observable degree without obtaining individual feedback. Despite accumulated research findings showing its ineffectiveness for learning (e.g. Truscott, 2010), the practice of giving individual written feedback has continued to be endorsed by teachers and students

alike—possibly due to the lack of compatible alternatives. Now that a new feedback approach has been proposed and its effectiveness demonstrated, this may be a time to reconsider appropriate and effective forms of feedback.

The answer to the second research question, although largely conditioned, is excitingly positive. Comparing the three drafts and three revisions, the only salient difference was found between the first and the third drafts; however, the significance here was robust across all five measures, despite the fact that one of the two raters disapproved of the argument rationale used by many students for the third task, which in turn led to decreased scores. It should also be noted that in between the three rounds of tasks, learners were engaged in the speaking component of the course and received no reinforcement on their previous written submissions, which make the result even more noteworthy. As noted from Figure 2, the iterative and ascending pattern of performance across the three tasks can be clearly observed. The only significant difference happens precisely where Truscott (2010) considered real learning could and should be identified in terms of writing performance; evidence of this had been missing until now. While it is acknowledged that learning is gradual and slow, a sign of this kind is encouraging and warrants more studies along this line of research.

For research question three, the data reveal that most students reported having learned about structure and organization. The things they claimed to do best at were generally consistent with what they reported having learned. Their difficulties, however, were mainly associated with a limited vocabulary and uncertainty about grammatically; the latter, as attested by some in the open-ended fourth question, pertained to the lack of individual feedback. These findings coincide with the pattern of sub-scores as shown in Table 2 and Figure 2, where lexis and grammar scores are comparatively lower. However, argument, organization, and holistic scores are all

higher. As for the degree of improvement from drafts to revisions, students advanced the most in terms of organization. These results imply that, based on the feedback lessons used in this study, learners improve 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 predominantly 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. On the other hand, lexis issues required a much longer period of time to work though as they were less subject to (at the mercy of) instruction, even though resources such as a thesaurus were introduced to students.

Researchers offer instructors a wide variety of interventions and facilitations that may help language students. However, teachers are often loaded with large classes and too much work, making it difficult to apply these tools. From a cost-effective standpoint, this new feedback method may serve as a viable alternative to the traditional method, which has been questioned for some time. The challenges for teachers associated with this proposed approach, however, are not less demanding, since this method involves a great deal of spontaneity and contingency. But the beauty of it lies in using teachers' time more wisely and strategically and making the complicated mission of communication more likely to succeed. The results are promising, but more careful examinations are needed in the future, such as a study design that separates the effects of individual written feedback and feedback lessons given to the group. When there is no additive effect in terms of the combination of the two types, then a comparison in terms of learning effectiveness and demands on the teacher can be investigated.

One point of caution has to do with the inherent nature of a feedback lesson for all. Instructional lessons targeted at an entire student group certainly do not cater to all

learner needs. When the class size is large 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 the whole-class lessons.

In addition, the feedback lessons were in a very primitive form. They served more as general directions than concrete procedures to follow. With its contingent nature, there probably would 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 Conclusion

Partly because of its exploratory nature, the study has several limitations. First, while pre-writing lessons and activities, such as brainstorming or providing lexical support, are common in L2 writing classes, they were absent from this study. The participating students started writing the drafts immediately after the prompts were presented. They worked independently until their first drafts were finished in a 30-minute timed writing. It is possible that the significant differences found between drafts and revisions would not have existed if some sort of pre-writing instructions had been offered. That said, giving pre-writing aides may very well have violated the principles this feedback instruction methodology is based on: as shown in Figure 1, learners felt empowered and involved from the beginning to investigate their current abilities, which led to greater focus during the follow-up activities and instruction.

Secondly, the kinds of revision students were rated on were actually a combination of the revision and editing steps. In Butler and Britt (2011), these two topics 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 deemed unnecessary. Future studies along this line may well take the different stages of revision and editing into consideration in terms of 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 about 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; however, students who did both tutorials showed no additive effect. While learners in the current study needed both kinds of instruction, it would be interesting to know if either of them could be more easily taught to students.

In conclusion, the current study proposes an alternative method of providing feedback to L2 writing learners. This new approach directs the teacher's effort and time away from giving individual written feedback on all drafts. Instead, the teacher diagnoses problems for the entire group and prepares a feedback lesson that focuses on the what and how of revision. In turn, the messages are more clearly communicated. The results are promising. Students' revised versions of the three tasks all improved significantly as compared to their drafts in terms of the five measures

employed. The effect also seemed to carry over from the first initial draft to the third one. 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. These results warrant further studies regarding alternatives to written corrective feedback.

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*(2), 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. (2007/8). Feedback that fits. *Educational Leadership, December 2007-January 2008,* 54-59.

Bruton, A. (2009). Improving accuracy is not the only reason for writing, and even if it were… *System, 37,* 600-613.

Butler, J. A., & Britt, M. A. (2011). Investigating instruction for improving revision of argumentative essays. *Written Communication, 28*(1), 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.

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.

Ewert, D. E. (2009). L2 writing conferences: Investigating teacher talk. *Journal of Second Language Writing, 18,* 251-269.

Ferris, D. R. (2003). *Response to student writing: Research implications for second*

language students. Mahwah, NJ: Lawrence 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). Cambridge: 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*(2), 207-234.

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

Hyland, K. (2000). ESL writers and feedback: Giving more autonomy to students. *Language Teaching Research, 4*(1), 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. (2011). Feedback revolution: What gets in the way? *ELT Journal, 65*(1), 1-12.

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

McGarrell, H., & Verbeem, J. (2007). Motivating revision of drafts through formative feedback. *ELT Journal, 61*(3), 228-236.

Min, H.-T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing, 15,* 118-141.

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.

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*(2), 199-218.

Parr, J. M., & Timperley, H. S. (2010). Feedback to writing, assessment for teaching and learning and student progress. *Assessing Writing, 15,* 68-85.

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*(3), 277-289.

Sadler, D. R. (2010). Beyond feedback: developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education, 35,* 5, 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.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46,* 327-369.

Truscott, J. (2010). Some thoughts on Anthony Bruton's critique of the correction debate. *System, 38,* 329-335.

Truscott, J., & Hsu, A.Y.-P. (2008). Error correction, revision, and learning. *Journal of Second Language Writing, 17,* 292-305.

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*(5),

519-533.

Appendices

Appendix A. Writing prompts and teaching points in the three revision lessons

	Task 1 (W4 draft + W5 revision)	Task 2 (W7 draft + W8 revision)	Task 3 (W12 draft + W13 revision)
Writing Prompts	Discuss the advantages and disadvantages of having a job while in college and then state your own opinion on this topic.	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.	In some universities, students do not choose their majors until the second or third year, while most students in Taiwan are put into different professional/academic departments before they are admitted. Which do you think may be a better arrangement for college students? Why? Provide detailed reasons for your answer.
Teaching Points in the Revision Lessons	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.	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.	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,

	3. Check if supporting details support the main points. 4. Analyze sentence structure and check grammar.	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.	such as subject-verb agreement, tense, and singular/plural forms.

Appendix B. Rubrics for instruction and grading

項目 Scoring Criterion/Level		**Excellent (90% & up)**	**Good (80%-89%)**	**Fair (70%-79%)**	**OK (60%-69%)**	**Poor (below 60%)**
		15, 14, 13	**12, 11, 10**	**9, 8, 7**	**6, 5, 4**	**3, 2, 1**
總分	**H**olistic Score 作文整體品質	極優秀	優良	良好	尚可	一般水準之下
內容 Content	**A**rgument 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)			
語言 Language	**L**exical 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)			

Navigation

Journal Home
Editorial Board
View Current Issue
View Future Issues
View Archives
View Contributing Authors
Guidelines for contributors
Review Board
Subscribe to Email List
Research Home
WMIER Home

English Teaching: Practice and Critique

Like a bell responding to a striker: Instruction contingent on assessment

Volume 11 Number 4 December 2012

Shu-Chen Huang (Foreign Language Center, National Chengchi University, Taipei, Taiwan)

This article is concerned pragmatically with how recent research findings in assessment for learning (AfL) can bring about higher quality learning in the day-to-day classroom. The first half of this paper reviews recent studies within Black and Wiliam's (2009) framework of formative assessment and looks for insights on how pedagogical procedures could be arranged to benefit from and resonate with research findings. In the second half, based on lessons drawn from the review, the findings were incorporated into an instructional design that is contingent on formative assessment. The concept of teacher contingency is elaborated and demonstrated to be central to the AfL pedagogy. Attempts were made to translate updated research findings into an English as a foreign language (EFL) writing instruction to illustrate how teachers may live up to promises offered by recent developments on AfL. This AfL lesson, situated in L2 writing revision, made instruction contingent on and more responsive to learner performance and learning needs. As shown in an end-of-semester survey, learner response to the usefulness of the instruction was generally quite positive.

pdf

English Teaching: Practice and Critique December, 2012, Volume 11, Number 4
http://education.waikato.ac.nz/research/files/etpc/files/2012v11n4art7.pdf pp. 99-119

Like a bell responding to a striker: Instruction contingent on assessment

SHU-CHEN HUANG
Foreign Language Center, National Chengchi University, Taipei, Taiwan

ABSTRACT: This article is concerned pragmatically with how recent research findings in assessment for learning (AfL) can bring about higher quality learning in the day-to-day classroom. The first half of this paper reviews recent studies within Black and Wiliam's (2009) framework of formative assessment and looks for insights on how pedagogical procedures could be arranged to benefit from and resonate with research findings. In the second half, based on lessons drawn from the review, the findings were incorporated into an instructional design that is contingent on formative assessment. The concept of teacher contingency is elaborated and demonstrated to be central to the AfL pedagogy. Attempts were made to translate updated research findings into an English as a foreign language (EFL) writing instruction to illustrate how teachers may live up to promises offered by recent developments on AfL. This AfL lesson, situated in L2 writing revision, made instruction contingent on and more responsive to learner performance and learning needs. As shown in an end-of-semester survey, learner response to the usefulness of the instruction was generally quite positive.

KEYWORDS: English teaching, EFL pedagogy, assessment for learning, L2 writing, instructional design.

> Respond properly to learners' enquiries, like a bell responding to a bell striker. If the strike is feeble, respond softly. If the strike is hard, respond loudly. Allow some leisure for the sound to linger and go afar. (*Record on the Subject of Education Xue Ji, Book of Rites Li Ji*, 202 BCE-220 CE)

INTRODUCTION

Modern educators believe that learning does not occur in a vacuum. For any individual learner, knowledge is co-constructed in the social-cultural context with "scaffolds" provided by more experienced others and peers. To facilitate such learning, Alexander (2006) describes "dialogic teaching," in which quality, dynamics and content of teacher talks are most important, regardless of institutional settings and classroom structures. Conventional IRF (initiation – response – feedback) turns are not adequate if the teacher's speech remains the core aspect and the learner's speech remains peripheral. In real dialogues, the learner's thinking and its rationale must be deliberately sought and addressed. Mercer (1995) describes this type of talk among teachers and learners as "the guided construction of knowledge." Right or wrong, learners' discussions of the subject matter provide an opportunity for self-reflection and self-assessment of their current knowledge. These discussions also allow the teacher to realize what needs to be taught. This understanding opens a path to deep learning. Mercer, Dawes, and Staarman (2009) provide an enlightening example of how dialogic teaching differs from more authoritative teacher talk and how this dialogue can be facilitated through pedagogical tools. Before explaining why the moon changes shape, these teachers designed "talking points" – a list of factually correct, controversial, or incorrect statements – and allowed pupils to discuss them

without fear of judgment. Based on these free and extended discussions, the teachers identified students' prior concepts, both right and wrong. This information improved follow-up teaching for the teacher and the pupils.

Perhaps a metaphor can help us to grasp the notion of dialogic teaching. We use the metaphor of treating an enquirer as a brass bell responding to a striker (see Figure 1). This metaphor comes from an ancient Chinese publication, *Record on the Subject of Education (Xue Ji),* a collection of ideas and conduct compiled by Confucian disciples. The 18th volume of the 45 volumes of the *Book of Rites (Li Ji)*, *Xue Ji* contains 20 sections with a total of 1229 Chinese characters. This classic's concise and archaic language must be translated into modern-day language and is subject to interpretation. Generally considered the earliest systematic documentation on education, it covers the purposes, systems, principles, and pedagogies of education. Many of its propositions remain true and inspiring after thousands of years, and a number of metaphors make its doctrines approachable. Among them, the bell metaphor illustrates the suggested attitude for teachers responding to students' questions. Teachers must be aware of and consider learners' capacity. Teachers are advised to assess learners' proficiency, to provide the appropriate amount of feedback at the right level, and to allow learners time to ponder and fully understand this feedback. These principles resonate remarkably with the convictions of dialogic teaching.

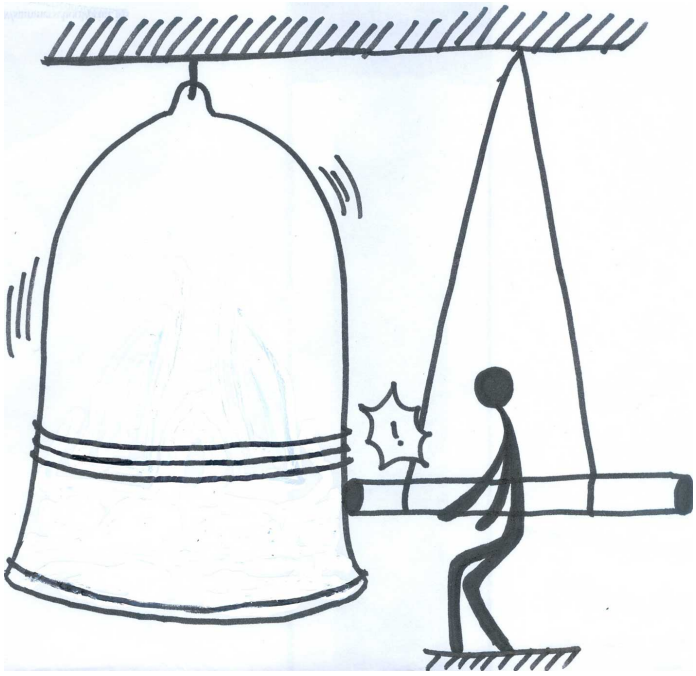


Figure 1. A bell responding to a striker

Studies on dialogic teaching, especially studies on the teaching of science and mathematics, have afforded illuminating examples of how effective dialogue helps learners to understand and, even more importantly, reveals misunderstandings. Hence, dialogue helps teachers to build on learners' existing knowledge and teach to their critical needs.

The current study on teaching English writing is slightly more complicated. First, it is not sufficient for learners to be able to talk about their understanding of writing; they

must perform. What is said well may not be performed correctly. Writing must be practised, and it is not practicable to wait until students have mastered every concept about writing. They may never be perfectly ready. In fact, students learn as they write. Furthermore, some basics of writing, such as coherence and unity, are abstract to learners, and understanding these concepts is usually a matter of degree rather than absolute knowledge. What may be the "talking points" (Mercer et al., 2009), as mentioned above, for a writing teacher? What kind of "answers" should a second-language writing teacher elicit to help her teach? The answer is straightforward: student writing. Student writing may disclose valuable information that allows a teacher to plan and structure her instruction. Yet, too often, student writing marks the end of a unit, and the teacher's written feedback returned with the students' writings are not used to their full advantage. As scholars have cautioned in other teaching contexts, "…the child's answer can never be the end of a learning exchange (as in many classrooms it all too readily tends to be) but its true centre of gravity" (Alexander, 2006, p. 25).

As Alexander (2006, p. 33) has noted, the ideas heralded by dialogic teaching are strikingly similar to ideas related to the assessment for learning presented by Black and his colleagues (Black, Harrison, Lee, Marshall & Wiliam, 2003). Because the focus of this article is a pragmatic application of dialogic teaching ideals in the second language writing classroom with a focus on teacher assessment and feedback on student writing, the following discussion will elaborate on recent studies in the area of learning assessment to justify the proposed approach to second/foreign language writing instruction.

ASSESSMENT FOR LEARNING

Formative assessment, in contrast to summative assessment which is high-stakes, standardised, evaluative, large scale and institutional, did not attract as much research attention as its counterpart did in the previous century. What educational roles it could play and how it was carried out were largely subject to classroom teachers' idiosyncratic discretion. The earliest systematic reviews are generally believed to be those of Crooks (1988) and Natriello (1987), focusing on the impact of evaluation practices on students. A decade later, Black and Wiliam (1998) used "the black box" as a metaphor to describe classroom assessment and started to explore the potential of revealing that black box, that is, using formative assessment for teaching and learning. A few research teams elaborated on possibilities of assessment *for* learning (AfL), as opposed to the more conventional role assumed for assessment, that is, assessment *of* learning (AoL). Earlier studies on formative assessment, or AfL, were mostly situated in science and math education at the school level, yet its influence has gradually expanded to other subject areas and institutional contexts.

L2 education has by no means been left out of this formative assessment movement. Although a few years ago concerns were voiced about the scarcity of such research in L2 classrooms (for example, Colby-Kelly & Turner, 2007; Rea-Dickins, 2004), the situation has been changing. Harlen and Winter (2004) depicted the development of formative assessment in science and math education in Britain for readers of the *Language Testing* journal. Six features of quality teacher assessment were identified, namely: 1) gathering and using information about learning processes and products; 2)

using indicators of progress; 3) questioning and listening; 4) feedback; 5) ensuring pupils understand the goals of learning; and 6) self- and peer-assessment. Many of these features have since been explored in other contexts. Leung (2004) located areas of challenge for AfL to be implemented in L2 classrooms, including conceptual clarification, infrastructural development, as well as teacher education. Cumming (2009), in a review of language assessment, pinpointed the difficulties in aligning curricula and tests and in describing or promoting optimal pedagogical practices and conditions for learning. More recently in 2009, *TESOL Quarterly* devoted a special issue on teacher-based assessment, offering a variety of perspectives on L2 formative assessment, in which teachers are empowered to make assessment decisions conducive to learning. In addition to general L2 issues, AfL was also introduced to and interpreted for specific subfields such as L2 writing (Lee, 2007a).

With the gradually widespread awareness and acceptance of formative assessment in different areas of education, the above-mentioned reviews helped scholars synthesise collective wisdom and attempted to lay down agenda for more researchers to follow, as well as principles for practitioners to apply. However, a genuine difficulty has gotten in the way, that being the lack of a unifying theory (Davison & Leung, 2009), one that could consolidate diffuse efforts and establish a future trajectory apart from the long-established, standardised testing paradigm.

An integrated AfL theory

Acting in response, Black, Harrison, Lee, Marshall and Wiliam (2003) summarised five types of activities based on evidence of effectiveness. These are sharing success criteria with learners, classroom questioning, comment-only marking, peer- and self-assessment, and formative use of summative tests. Subsequently, Black and Wiliam (2009) developed a two-dimensional framework to organise the various aspects of formative assessment. One dimension in their theory is the agent of learning which could be the teacher, a peer, or the learner him/herself. The other illustrates the stages of learning, these being the goal — "where the learner is going", the current status — "where the learner is right now", and the bridge between the two — "how to get there." Forged under Black and Wiliam's two dimensions, classroom learning based on formative assessment is believed to progress in the following temporal sequence (p. 8).

> *Where the learner is going*
> Step One: Clarifying learning intentions and criteria for success (teacher)
> Understanding and sharing learning intentions and criteria for success (peer)
> Understanding learning intentions and criteria for success (learner)
>
> *Where the learner is right now*
> Step Two: Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding (teacher)
>
> *How to get there*
> Step Three: Providing feedback that moves learners forward (teacher)
>
> *Where the learner is right now/How to get there*
> Step Four: Activating students as instructional resources for one another (peer)
> Step Five: Activating students as the owners of their own learning (learner)

Under the sequence depicted in the theory, it is clear that assessments serve as a backbone in interweaving a series of instructional activities in which teacher, peer, and learner are all actively involved. First, the teacher communicates learning goals and makes sure learners understand the goals through collaboration with peers in the classroom. Second, the teacher elicits and evaluates learner performance to make professional judgment on where the learner is relative to the stated learning goals. Thirdly, with goals and status quo acknowledged by both parties, the teacher then provides appropriate feedback to help learners go from where they are to where they aim to be. The first three steps form a cycle of assessment for learning across all three agents and three stages in the framework. It should also be noted that in the first step, the teacher is not the only agent, conveying the goals unilaterally. Deliberate emphasis is laid on learners to take on the goals through scaffolding from teacher and peers. That is also why the cycle does not stop at step three. After steps two and three, in which the teacher is the major acting agent with learners as the more passive recipients, peers and learners are then invited to mimic what the teacher has modelled in steps two and three, eventually taking over the responsibility of learning as that of their own.

This delineation of AfL instructional procedures is different in a number of ways from the conventional ones observed in most classrooms. First of all, in step one, instruction starts with deliberate clarification on learning intentions and criteria of success. Although these objectives are routine elements in syllabi and lesson plans, teachers usually do not spend much time communicating them to learners. Teaching what they think learners should and will learn is usually the primary part of instruction. On the other hand, the AfL step one, with a learner-centre crux, addresses learning goals more explicitly and brings students inside the "black box" of assessment and informs them with a sense of targeted criteria and standards for learning. Secondly, this communication is taken so seriously that learners are invited to, in addition to receiving information from the teacher, internalise the criteria through collaboration with peers. Finally, steps four and five further open up the "black box" to learners. These two steps are purposefully included in AfL after feedback is provided, which usually marks the end of a non-AfL instructional cycle. It is now generally acknowledged that learners do not automatically pick up messages in teacher feedback and improve (for example, Lee, 2007b, 2010; Price, Handley, Millar & O'Donovan, 2010). Such learning has to be meticulously planned and facilitated. In sum, the theory developed does not only serve research purposes, it also characterises an ideal AfL with the necessary instructional sequences. For a practitioner, the implications from these research developments are inspiring. What we need next is to translate these precepts into actual classroom actions.

A REVIEW OF THE LITERATURE

In this section, recent studies conducted under the framework of Black and Wiliam's (2009) formative assessment theory are reviewed. Although there is not always a fine line between each of the five steps of AfL instruction, it is believed that an examination of practice against the framework may help identify current developmens and inform practice. This stocktaking of the current development of different aspects of formative assessment serves as a basis for the subsequent design of an instructional unit that translates research findings into concrete classroom procedures.

Step One: Clarifying, sharing and understanding learning intentions and criteria for success (Teacher, Peer & Learner)

Research findings have made it clear that the communication of learning intentions and criteria for success is anything but straightforward (Price, O'Donovan & Rust, 2007). Learners are oftentimes not quite clear about these standards, for example, as revealed in Lee's (2011) case study of formative assessment in L2 writing (p. 101). Learners do not usually have a firm grasp of their learning goals, especially at the level teachers expect. To facilitate this communication, one tool commonly referred to is instructional rubrics, or rubrics used for teaching and learning, which usually contain both a list of criteria for evaluation and the gradation of quality for these criteria. Such rubrics help make teachers' expectations clear to learners. It is believed that even young school pupils can be taught to use them, or better even, to be involved in creating them with teachers (Andrade, 2000). Reddy and Andrade (2010) conducted a review of rubric use in higher education and found that: 1) it is generally well received by students but some teachers tend to resist changes to their normal practice; 2) it is associated with improved academic performance in some but not other studies; 3) simply making rubrics available to learners does not work. Students have to be deeply engaged, such as co-creating rubrics and using them for peer- and self-assessment; 4) the appropriateness of the language and content as well as the adaptation to the particular learner population with their characteristics considered is of paramount importance for rubrics to be successful. There are also cautions regarding increasing transparency of assessment criteria in higher education (Norton, 2004), with the fear that "grade-grubbing" students may end up strategically and tenaciously focusing on the superficial aspects of tasks and missing the genuine purpose of learning. It is, therefore, the teachers' responsibility to help learners eventually re-conceptualise assessment criteria as learning criteria.

Specifically, on the intricacy of articulation of assessment requirements for learners, O'Donovan, Price and Rust (2004) acknowledge its difficulty and side-effects, and give a comprehensive illustration of a spectrum of processes supporting the transfer or construction of learners' knowledge of assessment criteria. O'Donovan et al. argue that although the education community seems to be, in recent years, pursuing ever-increasing explicitness, mainly through verbal description, of assessment criteria in the hope that learners may better acquire that knowledge, they believe it is neither possible nor necessary. First, even in a small institutional context with a small group of experts teaching the same subject matter, standards are not definitive. The same set of evaluating criteria is subject to multiple interpretations by teaching staff, let alone learners. Moreover, knowledge is not always acquired through the explicit and precise articulation of information by others.

O'Donovan et al.'s conceptual framework, thus, features a continuum of explicit versus tacit knowledge transfer and the various channels of communication to be utilised both before and after the submission of student work. In the pre-submission stage, the more explicit means may consist of written learning outcomes, written marking criteria, tutor assessment briefings to students, and discussion with a tutor. Gradually moving toward the more tacit side are peer discussion, self-assessment, marking practice, and exemplars. After submission, the more explicit means for communicating standards are tutor-written feedback, peer-written feedback, and oral feedback and discussion with a tutor. The more tacit transfer channels include oral feedback and discussion with peers along with examples of excellent submitted work.

It is contended that no one single method is robust enough to convey the complicated message of standards and a pursuit for ever more explanatory details in assessment criteria is futile and meaningless. Following this line of analysis and argumentation, practical intervention suggestions are drawn, such as: 1) providing all relevant assessment information, that is, criteria and marking schemes, so that learners understand what is required of them in terms of basic and higher levels of performance; 2) engaging students in assessment activities so they may actively attend to the assessment information; and 3) providing opportunities for teacher-student exchange about the evaluation of performance both in the preparation stage and after teacher feedback on the assignment is received by students (Bloxham & West, 2007).

Step Two: Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding (Teacher)

Step two is for teachers to learn where students stand. This step, such as turning in a homework assignment, used to be the final step marking the completion of an instructional cycle and moving learners on to the next one. However, in an AfL sense, teachers have a lot to learn from learner performance. Together with its relevance to the criteria of success in step one, learner performance forms a basis to support the teacher's professional decision of what and how to teach. Without this knowledge, it is as if the teaching is shooting without having a target. This knowledge could be obtained both during teaching, such as through questioning learners and through other dynamic interactions in the classroom, and after teaching when the teacher receives and evaluates student performance on an assigned task.

Ways of eliciting evidence varies. It has to do with the nature of different disciplines. Some favour problem-solving, others oral presentations. Discussion of task-types may be more focused if it is kept within the same subject matter. Take L2 writing as an example, where the task is usually more straightforward. Learners write and produce a complete piece of writing according to prompts or guidelines given by the instructor. The prompts may range from more closed-ended ones, such as describing a series of frames in a comic, to more open-ended ones like having learners decide on their own topic within the limits of the given directions. This decision on task specification is usually based on learner proficiency and learning objectives. Task-type, as explicated by Torrance and Pryor (2001), could be more convergent or divergent, with the former testing *if* students have learned from the teacher's curriculum and the latter testing *what* they can do on their own curriculum. There are indications that task-type may influence students' motivation and self-regulated learning strategies (Huang, 2011).

Step Three: Providing feedback that moves learners forward (Teacher)

Research on feedback, compared to that of the other four steps in the AfL framework, is relatively prolific. Empirical studies (for example, Price et al., 2010; Wingate, 2010) on current feedback practice generally reveal situations where teachers are stressed out by the piles of homework to which they must give feedback, while students either do not care to read, do not understand, or, when they do, are frustrated by teacher feedback, creating a "no-win situation" (Lee, 2010, p. 46). Specifically in relation to L2 writing, Lee (2007b) investigated Hong Kong secondary classrooms and found that feedback was not exploited for AfL purposes. The common scenario, similarly, is quite often of a teacher meticulously circling all errors that can possibly

be found and of a student receiving his homework "awash in red" and losing confidence.

Pedagogical suggestions based on AfL philosophy provides some guidance for feedback practice. Brookhart (2007/8), in the context of teaching paragraph-writing to American fourth-graders, explicates the dual aspects of feedback, namely cognitive and motivational. First, feedback cannot move learners forward if it is beyond their cognitive capacity, or if they are not willing to pay attention to it, or if they do not believe in its usefulness. So it is important for teachers to adopt a "student's-eye view", rather than a "teacher's-eye view", in formulating feedback, taking into consideration learners' learning history, personal characteristics and current developmental level. Secondly, on the motivational aspect, well-intentioned feedback with information that may actually move learners forward cognitively may run the risk of being psychologically destructive if learners' self-efficacy is endangered, especially for low achievers. It is unrealistic and even damaging to expect students to fix all of the problems in a piece of work. Therefore, concludes Brookhart, teachers should inform learners what has been achieved, describe but not judge, be positive and specific, and guide students into the next step of improving their work.

Studies on feedback in higher education do not seem to be hugely different from those in primary education, although the complexity of knowledge in higher education is highlighted. Price et al., (2010) classify the feedback function into five categories, with each one building upon the previous type. The five categories are correction, reinforcement, forensic diagnosis, benchmarking and longitudinal development (feed-forward). They emphasise the difficulty of making all feedback explicit. For example, error identification is usually not straightforward in higher education. To reflect and respond to the multidimensionality of homework assignments in higher education, it is necessary to provide feedback that is partially gauged on explicated criteria and partially by more tacit professional judgment. The areas with potential for improvement in a piece of student work, if it is within the curriculum, could be pinpointed for revision. But if they are beyond the current curriculum, it may be more difficult and less necessary. Moreover, student responses also indicate that assessor-student dialogues and work samples are preferred over "tick box" feedback, the latter sometimes giving students a negative feeling and casting doubts on whether their work was indeed read. It is also cautioned that easily quantifiable indices may be very tempting, but the very essence of teaching and learning should still be the ultimate concern.

Similarly, in the context of higher education, the effect of feedback on academic writing has been demonstrated (Wingate, 2010). Students who responded to feedback did improve and did not receive the same old criticism for their revised work. But for low-achievers and less motivated learners, assessors should refrain from giving too many negative comments all at once or comparing learner work unfavourably with their more proficient peers. Quality and effectiveness of feedback need substantive cultivation in pedagogical content knowledge on the assessor's side. Parr and Timperley (2010) analysed feedback for student writing against AfL principles. They found that students of those teachers who were capable of providing quality feedback did improve more in writing than their peers whose teachers' feedback quality was lower. To bring feedback to the centre of the classroom, Hattie and Timperley (2007) put feedback and instruction on the two ends of a continuum. Sometimes instruction is

embedded in feedback and there is not such a fine line between the two. Leung (2007) also emphasizes the dynamic nature of assessment and discusses assessment *as* teaching. One unequivocal conclusion is that for feedback to work, it must be put it into use. This can be accomplished by affording timely opportunities and vigorous motivation for students to use teacher feedback in improving the quality of their work (Brookhart, 2007/8; Price et al., 2010).

Step Four: Activating students as instructional resources for one another (Peer)

In AfL, instruction is not yet complete after teachers have provided feedback, although it may be the case in many other instructional situations. Acquiring the ability to self-evaluate is an integral part of learning. And for that learning to happen, peers in the same classroom are valuable resources not to be left unexploited. Peer evaluation is a process often included in L2 writing, but its utility has been debated (Rollinson, 2005). One of the concerns which became the focus of some studies lies in the peers' ability that is yet to be developed. For example, Matsuno (2009) investigated this ability among Japanese university L2 writing students. Her findings indicated that peer-raters were more consistent and lenient than self-raters regardless of their own writing abilities. Bloxham and West (2007) claimed that simply giving written criteria and grade descriptors to peer raters is insufficient in making tacit assessment knowledge transparent to novice student raters. It has to be buttressed by marking exercises. They used peer assessment as a way to develop students' conceptions of assessment and identified its benefits, specifically in regard to the use of criteria, awareness of achievement, and ability to understand feedback. Price et al. (2007) found similar benefits of peer-review workshops from the viewpoint of both students and tutors in the usefulness of the process. Interestingly, Price et al. also discovered that, despite the positive perception, some students did not act upon the sound advice they received and got by with their unrevised work.

To bridge the gap between feedback given and feedback acted upon, Cartney (2010) used peer assessment as a vehicle and explored its influence in a case study. Learners, in groups of five, assessed and commented on each other's work. Besides the cognitive aspects usually discussed in literature, she found in peer assessment processes a deeply emotional nature. Students experienced anxiety or anger both in assessing others and in being assessed. Some were more reluctant to give true negative opinions, for fear of hurting feelings. Others felt that face-to-face communication was more efficient than exchanging on an e-platform. On the positive side, the peer process also stimulated in some a deeper self-reflection and initiated in others a more autonomous peer review practice on later work, where they had not been assigned to do so. In sum, the interpersonal and emotional consequences, part of the psychology of giving and receiving feedback, while generally not having been a focus in peer assessment studies, have to be taken into consideration if an AfL instructional design is to fully benefit from peer interactions.

Step Five: Activating students as the owners of their own learning (Learner)

The previous four steps are important building blocks for this final stage, where learners are expected to internalise the ability to self-assess performance against learning goals and, moreover, to use this capability to feed forward into independent future learning. To achieve this goal, after being supported by a scaffold from teachers and peers, learners try removing the teacher and peer scaffold to perform self-assessment. Research findings generally indicate a weak to moderate effect of

this practice. Chen (2008) trained Taiwanese college EFL learners to self-assess oral performance. By comparing their own scores with those of the teacher's, it was found that students improved significantly in rating accuracy after learning and discussion. Butler and Lee (2010) taught Korean sixth-graders to self-assess their English learning. They found positive but minimal effects on performance and confidence. Lew, Alwis and Schmidt (2010) investigated learners' self-assessment accuracy and their beliefs about its utility in higher education. Their learners' accuracy did not seem to improve over time and feedback did not seem to help either. Unlike results from Matsuno (2009) that peer-assessment accuracy was ability independent, as mentioned in the previous section on peer-assessment, Lew et al.'s findings suggest an ability effect on self-assessment. That is, learners with higher ability tend to self-assess more accurately. There was also no significant relation between this ability and the belief about the utility of self-assessment.

To understand the discrepancy between what teachers teach and what learners learn in carrying out formative assessment, Nicol and Macfarlane-Dick (2006) used a self-regulation model to analyse how a teacher-set task is translated into learners' own internal processes of goal-setting and strategy use, and how this learner internalisation may deviate from what the teacher has expected. Such discrepancy needs to be conceived more fully if healthy changes are to happen. Sadler (2010) points out that, in formative assessment, too much effort has been put on how teachers construct feedback, which may be off the mark. Self-assessment ability, as Sadler suggests, if cultivated only through teacher telling or perfectly-structured written feedback, cannot succeed, since there is a huge gap between the two parties' background knowledge and experiences in the target discipline, making the correct interpretation of complex messages almost impossible. There is a need to shift from the teacher's "disclosure" to the task's "visibility", that is, having students see and understand the reasons for quality. Three fundamental concepts are proposed for this visibility to be comprehended by young inexperienced learners. They are task compliance (to understand what is required of learners in the task), quality (to be experienced in making judgment), and criteria (to enable reasoning about the judgment). Sadler contended that, for this to happen, learner exposure to exemplars should be planned and not random. In addition, learners need to be given substantial evaluative experiences as a strategic part of the teaching design, rather than as the supplement it is in most current practices. By this intentional design on learner exposure and practice, together with the downplaying of teacher feedback, we may reasonably expect learners to be inducted into sufficient explicit and tacit knowledge to equip them with the ability to self-assess.

LESSONS LEARNED FROM AFL STUDIES

The above review within the framework of five steps in an AfL instructional sequence suggests that current teaching practices in general have some fundamental discrepancies from the AfL ideal. Traditionally, teachers unilaterally set teaching objectives, and from those objectives prepare materials and lesson plans before they have a chance to observe student performance, which is usually scheduled at the very end of one teaching cycle. One of the common ways to ensure learner-centeredness is to conduct a needs analysis. But usually a needs analysis encompasses the more general learner background information, such as learning needs and wants, rather than

a more specific diagnosis of where learners are in relation to the learning objectives. Furthermore, there is usually too fine a line between teacher instruction and learner performance, with instruction being central and performance peripheral. After well-planned instruction and delivery, homework guidelines are given and students are left alone to demonstrate their learning, usually outside of the classroom as an add-on to a lesson or mainly for record-keeping purposes. Assessment could, but usually does not, serve the purpose of learning.

Lee (2010) proposes a feedback revolution in L2 writing. Feedback is indeed central to AfL, because feedback has learner performance as a precedent focus, making learner performance an anchor point of instruction. But other than feedback, there is more to AfL than feedback alone. The AfL framework delineated above also emphasises communicating the criteria of success, and learners serving as instructional resources for each other and for themselves. All of these have a strong learner-centred orientation and are not adequately addressed in research and practice as compared to more narrowly defined "feedback". It is not just a feedback revolution that we need; it is an instruction revolution that puts feedback at the centre, where feedback and instruction are intertwined.

Lessons learned from the above literature review are strikingly similar to ancient Confucian wisdom as documented in the *Record on the Subject of Education (Xue Ji)* from *Book of Rites (Li Ji)* as early as 202 BC in the Han Dynasty. As the literal translation shows at the beginning of this article, the act of responding to learners is likened to a bell responding to a bell striker. The bell does not sound without anyone striking it. Depending on the power of the strike, the bell sounds in response to the relative power the striker shows. It makes no sense to sound loud if the striker is feeble, because the striker is not able to take it in; simply circling all errors would not benefit the learner. The "two stars and one wish" rule of thumb in providing feedback has the same philosophy – giving learners the confidence to go on and giving them just enough to absorb and digest. Finding the right balance for what to give is crucial on the responder's side if the striker is taken as central. Moreover, the bell should be patient enough to allow time and leisure for the sound to linger and go afar. Learners also need time to process what has been taught and to internalise the learning.

Lee's (2007b) recent investigation of L2 writing instruction in Hong Kong revealed a classroom that predominantly used assessment *of*, but not *for*, learning. Teachers spent lots of time making comprehensive error corrections as feedback and students were discouraged with such feedback. Learner dissatisfaction was apparent. Implications drawn are overwhelmingly in accordance with the learner-centred AfL explicated above. First, learners demonstrate a wish to learn more about the criteria of "good" writing. They want teachers to show more good examples so they know what to look up to. This is step one in Black and Wiliam's framework. Second, for the content of feedback, learners expect pervasive error patterns pointed out rather than all errors. This takes learners learning capacity into consideration. Using the bell-striker metaphor, sounding loud would only overwhelm a feeble striker. Instead, the teacher should sound just enough so the striker can take it and get ready for the next firmer strike. This connects to another point raised by learners in Lee's study, that is, giving them a chance to improve their writing. This is key to the success of feedback as already made clear in many feedback studies. Finally, in order for learners to learn from feedback, learners indicated that written feedback alone is not very useful.

Teachers' oral feedback given to the whole class is more useful and there is a need for more in-class discussion about the writing. The current peripheral outside conferencing should be moved into the centre stage of the classroom.

The above discussion illustrates a concept mentioned, but much less frequently elaborated in AfL research. That is the concept of "contingency" (Black & Wiliam, 2009), in which a teacher acts upon learner performance. As pointed out by Black and Wiliam (2009), "[T]hese moments of contingency can be synchronous or asynchronous. Examples of synchronous moments include teachers' 'real-time' adjustments during one-on-one teaching or whole class discussion. Asynchronous examples include teachers' feedback through grading practices, and the use of evidence derived from homework…to plan a subsequent lesson" (p. 10). Such lessons are subject to all kinds of variation and make a model lesson almost impossible. Even so, the author attempted in the following L2 writing lesson to translate summarised AfL principles into concrete classroom procedures.

A CONTINGENT L2 WRITING LESSON AND LEARNER RESPONSE

The following AfL lesson plan was implemented in the Fall 2011 semester as part of an integrated-skill Freshman English course in a university in Taiwan. It was the author's initial effort in a contingent instruction plan, by redirecting effort, spending less time on areas where learners would not benefit and more time on areas that may facilitate AfL. In this plan, the following objectives were targeted in order to meet the AfL principles discussed above.

First, give learners enough guidance and practice to decipher the criteria of good work, by providing exemplars of various gradations and ample time for discussion. As Davison and Leung (2009) assert about teacher-based assessment, "…trustworthiness comes more from the process of expressing disagreements, justifying opinions, and so on than from absolute agreement" (p. 409).

Second, move the major part of teacher instruction from before learner performance to after it. Teaching without learner performance as a reference point is like sounding the bell without a striker. Effective teaching is an act contingent upon learners' lacks and needs (Black & Wiliam, 2009). Situate the major part of teacher preparation between the time when learner work is collected and when it is returned for revision on the learners' side. This is, in fact, much more challenging for the teacher in that the lesson plan incorporates knowledge of where learners are and where they are going.

Third, devote more class time to feedback discussion, as strategically planned and not random. Instead of writing feedback for each individual without follow-up elaboration, the teacher organises feedback from the pervasive patterns found in learner performance and brings it to class for face-to-face discussion with the learner group.

Based on the above literature review and student learning history, a mini instructional program catering to principles of AfL was designed. This AfL writing program has an opinion essay as the target genre and expects learners, by the end of the instruction, to be able to write a coherent multi-paragraph essay of at least 300 words discussing the

pros and cons of an issue and expressing clear personal opinions with adequate supporting details. Major procedures are described in Table 1. How each step is to be carried out and the rationale behind it are discussed below.

Time	Procedures	Main agents	AfL steps & principles	Instructional materials
Unit 1	1. Retrospective reflection on past L2 writing experiences 2. Introduction to a new genre – argumentation, launched on TOEIC opinion essay and its official grading rubrics 3. Collective rating on TOEIC samplers using TOEIC rubrics, summarising results and whole-class discussion	• Teacher • Learner group	• Gathering information • Formative use of summative tests	• Samplers • Description of criteria and standards
Unit 2	1. Timed writing, first draft, 30 minutes 2. Breaking down the holistic rating from the previous unit – introducing the customised instructional rubric 3. Demonstrating using the rubric against a TA sampler 4. Group peer review workshop practicing the use of rubric and providing comments (2 stars and 1 wish)	• Individual learner • Teacher • Peer	• Ensuring understanding of the goal • Peer assessment	• TA sampler • Instructional rubric • Peer work
Between Units 2 and 3	Teacher reviewing learner work, inducing common patterns and deciding on instructional points, selecting samples to generate problem sets for instruction, preparing for a demonstration of	• Teacher	• Gathering information • Preparing for a contingent lesson	• Learner work

	how writing can be improved through various strategies. No grading of each individual work.			
Unit 3	1. Returning drafts for self-assessment and reflection 2. Delivering the contingent lesson, inviting learner discussion and problem solving, questioning to ensure understanding 3. Allowing time for reflection and discussion before revision 4. Individual revision 5. Self-assessment of the revision against first draft, personal reflection on the overall writing experience	• Teacher • Learner group • Individual learner	• Providing feedback • Questioning and listening • Ensuring understanding	• Excerpts of learner work • Instructional rubric
Units 4 and 5	The second round repeating procedures from Unit 2 through Unit 3; altering teaching points and building upon previous learning			
Units 6 and 7	The third round repeating procedures from Unit 2 through Unit 3; altering teaching points and building upon previous learning			

Table 1. An AfL instructional plan for L2 writing revision

To tide learners over from their previous EFL writing experiences, the teacher first probed learners to reflect upon their past writing assignments in a whole-class discussion. Points of discussion included the type of writing prompts, the length of time given, the length of writing in terms of number of words expected of them, how they prepared to write, what they did during writing, the type of feedback they got from teachers, what they did after getting feedback, and what was considered to constitute a good piece of writing. This discussion brought learner awareness to the

surface and at the same time provided the instructor important information for future planning of writing lessons contingent on learner needs.

At the beginning, the target genre (in this case, an opinion essay) and the difference between it and the learners' past writing genres was introduced. For the convenience of communication, the TOEIC writing component, and specifically its standard question type eight, an opinion essay, was introduced. To prepare learners for an assessment-sensitive writing lesson, success criteria had to be deliberately communicated by the instructor and clearly felt by the learners. According to O'Donovan et al. (2004), both explicit verbal descriptions of criteria and the more tacit knowledge of what constitutes a good piece of work should be taught through various information channels. In order to do that, the author/instructor, by gauging the writing tasks on the TOEIC opinion essay, first introduced the official scoring criteria and the various standards ranging from five (the highest) to zero (the lowest) as released by the examination institution on its website (Educational Testing Service, 2011). The verbal descriptions provided by ETS, encompassing content, organisation, lexical usage and grammar, were explained to students by the instructor. Another useful resource was a set of five examinee samplers matching a range of high to low scores against the official criteria (Trew, 2006). Instead of showing the given scores directly along with the samplers, the instructor engaged students in a collective rating exercise. A sample piece was shown to students first. They were given a few minutes to read and evaluate the work and assign a score to it individually without discussion. Once they were ready, the instructor asked for a show of hands and counted the class result on the blackboard. Learners were then invited to justify their choices and discuss their disagreement. After discussion, the actual score assigned in the TOEIC preparatory text, together with its explanation, was revealed to students. This whole-class exercise was repeated five times so learners were exposed to a variety of performance standards under the same writing prompt and the same scoring scale. The discrepancy among student raters and the collective tendency of rater scores were also highlighted to point to them the nature of qualitative rating. It was hoped that these introductory procedures prepared learners well for their own writing by understanding the criteria and feeling a sense of ownership for their assessment capability.

Before Unit 2, three documents had been prepared for use in class. First, the instructor translated the examiner-scale used in the previous week into an instructional rubric that was expected to better serve instructional purposes (Andrade, 2000). In addition to the holistic score, four sub-scores on argument, organisation, lexical use, and grammar were added. Each of the five-scale levels were mapped to a 100-point scale, which learners were more familiar with from their past school experiences, and further divided into three finer levels. Verbal descriptions of the four subcategories were written concisely in learners' first language in the hope that learners could understand easily. The second document was a sample essay written by the course teaching assistant who was at that time a senior student at the same university and whose English ability was comparatively high among the university's entire student population. She was told to write under the same prompt but was not given any instruction. It was hoped that her writing could be given to learners once they had completed their own, to serve both as a high-standard sample and, since it would not be a perfect piece of work, a reference on how the instructional rubric could be used in assessment for learning. The third document was an assessment table to be used in

peer assessment workshops. The table required learners to write down the name of the author, the name of the rater, a holistic score ranging from 15 to 1, four sub-scores ranging from 15 to 1, two positive comments, and one constructive comment on what needed to be improved.

In the second class meeting, learners first wrote individually for thirty minutes. After the first draft, the instructor introduced and explained the instructional rubric. The TA writing sample was at this point distributed so learners could see how the same writing prompt was responded to differently by a more proficient peer writer. The instructor, after giving a few minutes of silent reading time to the class, then used the sample to demonstrate ways of using the instructional rubric and giving comments. More specific to what existed in this TA sample, the teacher illustrated how the writing responded to the prompt nicely in its organisation and argumentation. Among the three points used to argue the writer's position, it was pointed out to the class that the first point was well supported and developed. In contrast, the second point manifested a typical example of lack of a meaningful connection between sentences and consequently it became less coherent. Other minor sentential errors were also elaborated and explained. By way of the teacher's analysis of the sample work, learners observed a good sample and learned why it was good. They also noticed what was not so good and could be improved. After this demonstration of rating and giving comments, learners were then put into groups of three or four for peer review. Each student received three rating table slips and their first drafts were rotated among group members. Each student was required to read two to three peer drafts and write down rating scores and comments. The student author then collected rating tables for his/her work from other group members, read and clipped the filled forms on top of his/her writing paper, and submitted the work to the teacher.

The instructor's actual writing lesson had not been planned prior to this moment. The teacher's time was not spent on grading and commenting on each individual piece of work, since past studies have shown that effort spent in doing so is not so effective. Rather, student work was reviewed in order to find common and pervasive problem patterns. After teaching points had been identified, learner excerpts were selected and areas in need of improvement were highlighted. These materials were designed into problem sets for the next class session so as to probe learners to tackle the problem and to foreground for the teacher her instruction on revision strategies.

In the third class meeting, the first drafts were returned to the learners. They reviewed peer comment sheets again and performed self-assessment against the same instructional rubric – a way to refresh their memory from the previous week and to connect the learned lesson to the new one. Following self-assessment, the rest of the first session was an EFL writing revision instruction unit contingent on learner performance exhibited in their first draft. An important point is that areas for improvement were always presented to students in the form of a problem set. The teacher allowed ample time for learner groups to ponder the problem and discuss possible answers and solutions. After group discussions, in which learners clarified and consulted one another's opinions, their ideas were then elicited and challenged by the teacher in whole-class discussion. It was at this point that the teacher demonstrated how a seasoned writer tackled the same problem with better, established strategies. The instructor eventually summarised the discussion into a few practical strategies for learners to apply in the next period of revision. Comprehensive coverage

of all problems identified was not the aim of the instructional content of this session. How much learners could take up was more important than how much there was to be taught.

The follow-up revision session was a time when learners worked individually with all resources nearby. Consultation of dictionaries, small-group discussions, and asking for help from the instructor and TA were all encouraged. When the time for revision was coming to an end, learners were asked to perform a final self-assessment on their revised version. To round off this writing and revision experience, they were asked to reflect and record strategies applied. This writing and revision cycle was expected to be repeated a few more times for more practice. Hopefully, each practice would feed forward to the subsequent cycles of writing exercises.

The contingent EFL writing revision lesson illustrated above is believed to refocus an instructor's effort from the laborious and seemingly ineffective individual paper marking to a holistic analysis of the bigger picture of common areas in need of teacher instruction. Instead of trying to correct each and every particular learner mistake, the instructor put problems into a few manageable entries and used learner excerpts as a point of departure for tailor-made instruction. Moreover, learners were invited to try problem-solving, group discussion and articulating their ideas, before the teacher offered her strategies. This step in engaging learners provided a chance to activate learner knowledge and awareness and, at the same time, gave the teacher a chance to know what learners actually could and could not do. This is what makes a contingent instructional lesson, one that derives from AfL principles, stand out from other instructional programs.

At the end of the semester, a questionnaire was posted on the course e-platform Moodle to invite anonymous learner feedback on the usefulness of the various components of this AfL writing instruction. The questionnaire listed 16 items of materials or exercises chronologically and asked learners to comment on the usefulness of each item in improving their English writing ability. The close-end choices were "very useful", "somewhat useful", "not quite useful", and "not at all useful." Among the 107 students enrolled in three sections taught by the author researcher, 61 participated in this online survey. Results were tallied and shown in Table 2. Learners were generally quite positive about this learning experience as the vast majority indicated the materials and tasks as very helpful or somewhat helpful. No one referred to any of the 16 items as "not at all useful". Means of each item were calculated and listed on the rightmost column. In particular, learners seemed to rate instruction and facilitation guided by the teacher much more positively than interactions with their peers. Other than items involving the teacher, TA model essays, revising learners' own first drafts, and writing the first drafts were rated at around 3.50 out of a possible maximum of 4. This information on learners' perception of the usefulness of various components in the contingent writing lesson could be a practical reference in refining future courses aiming to promote assessment for learning for a similar learner population.

How useful was each of these items in improving your English writing ability?	Very useful		Somewhat useful		Not quite useful		Not at all useful	Mean[1]
1. TOEIC essay criteria and descriptors	15	25%	43	69%	4	7%	0	3.11
2. Rating TOEIC essay samplers	21	34%	36	59%	4	7%	0	3.21
3. Knowing my peers' rating of TOEIC samplers	11	18%	42	69%	8	13%	0	2.92
4. The instructional rubrics and score sheets	32	52%	28	46%	1	2%	0	3.49
5. Writing the first draft	30	49%	28	46%	3	5%	0	3.39
6. TA model essays	33	54%	27	44%	1	2%	0	3.51
7. Peer review workshops	19	31%	39	64%	3	5%	0	3.21
8. Scores on my essays given by peers	8	13%	42	69%	11	18%	0	2.77
9. Comments on my essays given by peers	14	23%	40	66%	7	11%	0	3.00
10. Teacher's mini-lectures on revision	42	69%	19	31%	0	0%	0	3.69
11. Examples used in teacher's mini-lectures	41	67%	20	33%	0	0%	0	3.67
12. Teacher's demonstrations of revision	32	52%	29	48%	0	0%	0	3.52
13. Writing resources introduced by the teacher	29	48%	28	46%	4	7%	0	3.34
14. Revision checklist	12	20%	38	62%	11	18%	0	2.84
15. Revising my own drafts	39	64%	18	30%	4	7%	0	3.51
16. Selected peer sample essays	22	36%	38	62%	1	2%	0	3.33

[1] Calculated by assigning 4, 3, 2, and 1 to each response of "very useful", "somewhat useful", "not quite useful", and "not at all useful" respectively.

Table 2. Results of the end-of-term survey

CONCLUSION

The concept of teacher contingency demands a great deal more from teachers, for they need to have a much greater capacity than those who teach with a pre-determined syllabus and from a well-structured textbook. The teacher needs to know his/her student population well, understand their past learning history and beliefs, diagnose their difficulties, induce from learner work common patterns so as to prioritise teaching points, probe learners to think and elaborate, and make real-time decisions to provide useful guidance. It was, however, not such an overwhelming mission. As the sample L2 writing revision lesson has shown, what learners may learn is not so

foreign to an experienced teacher. The thrust of such instruction lies in tying instructional content to learners, including learner sample work and learner's demonstrated capability. It makes the teaching directly relevant to student learning. This is what assessment for learning is trying to achieve.

As quoted at the beginning of this paper, ancient Confucian wisdom addressed how a teacher could best respond to learner inquiries. The responder does not provide all the knowledge he/she has on the subject, since doing so would run the risk of overwhelming and discouraging the inquirer. Like a bell responding to a bell striker, the teacher takes the strength of the strike into consideration, gives just enough so that the learner can take it in, allowing him or her time and leisure to ponder on the response so that the sound may linger and go afar.

ACKNOWLEDGEMENT

This work was supported by the National Science Council, Taiwan. [grant number NSC100-2410-H-004-186-MY2]

REFERENCES

Alexander, R. (2006). *Towards dialogic teaching: Rethinking classroom talk.* (3rd ed.) Cambridge, England: Dialogos.

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership, 57*(5)*,* 13-18.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice.* Buckingham, England: Open University Press.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability, 21*(1), 5-31.

Bloxham, S., & West, A. (2007). Learning to write in higher education: Students' perceptions of an intervention in developing understanding of assessment criteria. *Teaching in Higher Education, 12*(1), 77-89.

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

Butler, Y. G., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing, 27*(1), 5-31.

Cartney, P. (2010). Exploring the use of peer assessment as a vehicle for closing the gap between feedback given and feedback used. *Assessment and Evaluation in Higher Education, 35*(3), 551-564.

Chen, Y.-M. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research, 12*(2), 235-262.

Colby-Kelly, C., & Turner, C. E. (2007). AFL research in the L2 classroom and evidence of usefulness: Taking formative assessment to the next level. *The Canadian Modern Language Review, 64*(1), 9-38.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438-481.

Cumming, A. (2009). Language assessment in education: Tests, curricula, and teaching. *Annual Review of Applied Linguistics, 29,* 90-100.

Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly, 43*(3), 393-415.

Educational Testing Service. (n.d.) About the TOEIC® speaking and writing tests. Retrieved October 12, 2011 from http://www.ets.org/toeic/speaking_writing/about.

Harlen, W., & Winter, J. (2004). The development of assessment for learning: Learning from the case of science and mathematics. *Language Testing, 21*(3), 390-408.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Huang, S.-C. (2011). Convergent vs. divergent assessment: Impact on college EFL students' motivation and self-regulated learning strategies. *Language Testing, 28*(2), 251-271.

Lee, I. (2007a). Assessment for learning: Integrating assessment, teaching, and learning in the ESL/EFL writing classroom. *The Canadian Modern Language Review, 64*(1), 199-213.

Lee, I. (2007b). Feedback in Hong Kong secondary writing classrooms: Assessment for learning or assessment of learning? *Assessing Writing, 12*(3), 180-198.

Lee, I. (2010). What about a feedback revolution in the writing classroom? *Modern English Teacher, 19*(2), 46-49.

Lee, I. (2011). Formative assessment in EFL writing: An exploratory case study. *Changing English, 18*(1), 99-111.

Leung, C. (2004). Developing formative teacher assessment: Knowledge, practice, and change. *Language Assessment Quarterly, 1*(1), 19-41.

Leung, C. (2007). Dynamic assessment: Assessment for and as teaching? *Language Assessment Quarterly, 4,* 257-278.

Lew, M. D. N., Alwis, W. A. M., & Schmidt, H. G. (2010). Accuracy of students' self-assessment and their beliefs about its utility. *Assessment and Evaluation in Higher Education, 35*(2), 135-156.

Matsuno, S. (2009). Self-, peer-, and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing, 26*(1), 75-100.

Mercer, N. (1995). *The guided construction of knowledge: Talk amongst teachers and learners.* Clevedon, England: Multilingual Matters.

Mercer, N., Dawes, L., & Staarman, J. K. (2009). Dialogic teaching in the primary science classroom. *Language and Education, 23*(4), 353-369.

Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist, 22*(2), 155-175.

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*(2), 199-218.

Norton, L. (2004). Using assessment criteria as learning criteria: A case study in psychology. *Assessment and Evaluation in Higher Education, 29*(6), 687-702.

O'Donovan, B., Price, M., & Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education, 9*(3), 325-335.

Parr, J. M., & Timperley, H. S. (2010). Feedback to writing, assessment for teaching and learning and student progress. *Assessing Writing, 15*(2), 68-85.

Price, M., Handley, K., Millar, J., & O-Donovan, B. (2010). Feedback: All that effort, but what is the effect? *Assessment and Evaluation in Higher Education, 35*(3), 277-289.

Price, M., O'Donovan, B., & Rust, C. (2007). Putting a social-constructivist assessment process model into practice: Building the feedback loop into the assessment process through peer review. *Innovations in Education and Teaching International, 44*(2), 143-152.

Rea-Dickins, P. (2004). Editorial: Understanding teachers as agents of assessment. *Language Testing, 21*(3), 249-258.

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment and Evaluation in Higher Education, 35*(4), 435-448.

Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT Journal, 59*(1), 23-30.

Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment and Evaluation in Higher Education, 35*(5), 535-550.

Torrance, H., & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal, 27*(5), 615-631.

Trew, G. (2006). *Tactics for TOEIC – Speaking and writing tests.* Oxford, England: Oxford University Press.

Wingate, U. (2010). The impact of formative feedback on the development of academic writing. *Assessment and Evaluation in Higher Education, 35*(5), 519-533.

Manuscript received: March 13, 2012
Revision received: July 10, 2012
Accepted: January 8, 2013

附件五

國科會補助專題研究計畫項下赴國外(或大陸地區)出差或研習心得報告

日期：100 年 11 月 21 日

計畫編號	NSC100－2410－H－004－186－MY2		
計畫名稱	學習動機、課室評量與教師權變		
出國人員 姓名	黃淑真	服務機構 及職稱	政治大學外文中心副教授
出國時間	101 年 11 月 11 日 至 101 年 11 月 18 日	出國地點	日本岡山 岡山理科大學、岡山國立大學

一、 國外研究過程

本次研究為配合計畫內容，探討日本地區大學課室評量之做法與教師觀點，主要連絡人員為岡山理科大學（Okayama University of Science）英語教師西岡啟治先生，扣除出發與回程，六天中共進行下列研究活動：

1. 對相關英語教師做學術演講

為導入研究議題相關之討論，並應安排行程之西岡先生要求，此行首先安排個人對大學英語教師之學術演講，講題為 “English Education in National Chengchi University, Taiwan”，內容報導台灣的大學英語教育現況，涵蓋大學學科能力測驗與指定科目考試中的英語考試內容、目前各大學英語教育現況等概述、政大的英語課程架構及課程指標、個人當期課程舉隅、英語畢業門檻制定的現況與背後的社會成因、以及其它校內英語學習相關活動等。與會者對台灣的大學英語教育均留下了深刻的印象，

並於會後提出許多問題，表達了深入瞭解的興趣。

當天在場教師多為在日本大學教授英語多年的專兼任教師，對台灣現象與日本的差異深感興趣。藉由這個機會，我認識了多位同行，成為他們對台灣教育瞭解的窗口，也開啟了未來進一步交流的可能。幾個未來連繫的可能對象為國際交流局局長西野雅二教授、負責學習支援中心的西川憲一先生、財團法人岡山日米文化協會的副理事長平松進先生等。

2. 拜訪國立岡山大學言語教育中心

經岡山先生協助，安排了一場禮貌性的拜會，接待我的有言語教育中心的主任亦為副校長的阿部宏史先生，言語中心的副主任劍持淑教授，及言語中心的副教授荻野勝先生等共四人，進行三十分鐘的討論。岡山大學準備了相當詳細的文宣資料，介紹學校的規模及言語教育中心的任務，包含中心業務與課程架構等，也表達了未來擴大合作的意願。在他們的國際合作名單上，目前台灣的簽約學校僅有台大與台北科大兩所，仍有相當的擴展空間。

3. 岡山大學 English Café

岡大的劍持淑教授等人大力推薦我體驗他們的 English Café，名稱聽來不陌生，因為我們也舉辦 International Culture Café ，但兩邊的運作上有許多不同處。這是一個常設性的地點，目地在鼓勵本地學生與外籍學生到這個溫馨的場地來認識新朋友，並以英文做為溝通的語言。它的位置在一間校園咖啡廳的隔壁，大約有一般 35 人教室的兩三倍大，牆上貼滿了各式各樣的活動海報及語言學習相關訊息，另有一個大螢幕，當時

正以較小的音量撥放 CNN 英語新聞，有許多小型討論桌，整個空間充滿了熱鬧也有點混亂的氣氛。我在一個午飯後的時間前往，當時有不少學生圍成小桌，輕鬆地聊天或做作業，令人感覺到校園的年輕與歡樂。

負責 English Café 的是外國語教育中心的宇塚万里子副教授，她與我聊了相當多的實際經營經驗。原本設立 English Café 也是希望能提供學生實際練習英文的機會，但後來發現許多外籍生到 Café 來，是希望在其許多使用日文的課業上能得到本地生的幫忙，雖與設立初衷略有出入，但仍不離國際交流的宗旨，因此也就不予強制要求在 Café 使用的語言。漸漸地，來到 Café 的學生愈形多元，使用的語言也五花八門。

我也訪問了當時在現場的學生，他們都非常活潑、願意表達意見，有的小桌一邊在做功課的同時，一邊還一同在桌上玩著疊疊樂的積木，得知我從台灣來，說的是中文，立即也找到了一位大陸青島去的學生來跟我聊聊。學生們表示日文相關的功課很多，能有個 Café 固定的場所，讓外籍學生向本地生求助是他們最大的需求，此外在輕鬆無拘束的場地與氣氛下結交各國朋友也是他們常到 English Café 的原因。

4. 課室觀察

此行經西岡教授的協助與陪同，進行一週的課室觀察，均是西岡先生的大一英文課。

西岡先生曾於五年前至台灣進行學術交流，經友人介紹，找到當時正在交通大學語言中心暨英語教學研究所任教的筆者，並到新竹停留數日，進行課室觀察，分別到筆者的大一英語會話與英語聽力課程中觀察課程進行實

況，他對台灣大學生的活潑和勇於課堂發言留下深刻印象，並提出了許多問題，我也趁便請他以照片對學生說明日本大學生的校園生活與外語學習情形，學生對於能獲知他國同儕的大學生活與學習情形，且一切以英語溝通，感到十分興味盎然。當時我因公務繁忙，課外時間遂安排班級中的學生進行簡單的介紹與招待，獲得學生普遍的支持，包括有兩個小組的同學進行校園導覽及帶領西岡先生體驗附近的清大夜市。西岡先生的老派日本紳士風度，很受學生喜愛，學生也跟西岡先生建立了忘年之交，許多人因這次以英語進行跨文化交流經驗的啟發，後來有很強的動機申請於大學期間出國做交換學生，有三位也與西岡先生保持了相當一段時間的連繫。

很高興五年後還能與西岡先生保持連繫，並進行這次的訪問研究。西岡先生的學術訓練於岡山大學完成，已在岡山理科大學任教英文科十七年，藉由他的幫忙，我得以在這兩所大學進行研究。

當週於西岡先生的課程中觀察，並與學生互動，實地了解課程進行實況，包括課本的使用（與台灣相似）、班級的人數、師生互動的情形，以及最重要的，教師在教學現場的即時評量與因應等現象。我所見到的這批日本學生，似乎在學習能力與動機上都較接近本地的技職校院學生，老師的教法也較少活潑的安排或互動，而是以講授或翻譯為主體。

幾次課堂上還有另一位觀察者，是來自美國俄亥俄州的一位年輕女性實習教師，她在岡山大學預計停留一年，參與所有岡山先生的課程，有時亦負責上課。對於當地學生有這樣由母語老師的機會

二、 研究成果

此行中除延申的訪問與拜會之外，最主要的是完成了一些課室觀察筆記及訪談錄音，也藉由實地在現場感受的日本大學英語教育的脈動。所得的第一手資料還有待仔細整理、轉錄成逐字稿，並做詳細的閱讀與分析。

三、 建議

1. 本校對外宣傳的文宣品應可製作更為精美並廣發給師生交流使用

2. 對於目前我在外文中心負責的外語自學中心,此行得到了一些改善經營方式的啟發。過去強調視聽設備的自學中心，以提供影音語言學習設備為主，學生多半到此來聽錄音或看影片，以期精進語言能力。一方面經費有限，設備與影音軟體難以大幅擴充，另方面這樣的設置已漸漸因時代與科技的改變失去了擴充的必要。如今學生在家中透過電腦與網路，都可輕易取得這些影視資料，自學中心很可能逐漸變成蚊子館。我們雖有 International Culture Café 的活動且受到好評，但畢竟是一次性的，舉辦完學生就各自離開，直到下次再辦活動，才有再交流的機會。因此，將自學中心改成一個熱鬧輕鬆的國際交流場所，似乎更符合現今的狀況，我會在中心主任工作結束前好好思考規畫這個可能。

四、 其他

攜回資料有下

- 所觀察課室之課堂講義兩份
- 2012 岡山大學概要、岡山大學 2013 案內、Okayama University Brochure、English Café brochure、岡山大學言語教育中心 brochure

附件五

國科會補助專題研究計畫項下赴國外(或大陸地區)出差或研習心得報告

日期：102 年 2 月 20 日

計畫編號	NSC100－2410－H－004－186－MY2		
計畫名稱	學習動機、課室評量與教師權變		
出國人員 姓名	黃淑真	服務機構 及職稱	政治大學外文中心副教授
出國時間	102 年 1 月 22 日 至 102 年 1 月 30 日	出國地點	美國 Minnesota St. Cloud State University

一、 國外研究過程

本次研究為配合計畫內容，探討美國大學課室評量之做法與教師觀點，主要研究地點與對象為位於明尼蘇達州的 St. Cloud State University，扣除出發與回程，七天中共進行下列研究活動：

1. 對 TESL Program 英語教師做學術演講

為導入研究議題相關之討論，並應 St. Cloud State University TESL Program Director – Professor James Robinson 要求，此行安排個人對大學英語教師之學術演講，講題為 “English Education in National Chengchi University, Taiwan”，報導台灣的大學英語教育現況，涵蓋大學學科能力測驗與指定科目考試中的英語考試內容、目前各大學英語教育現況等概述、政大的英語課程架構及課程指標、個人當期課程舉隅、英語畢業門檻制定的現況與背後的社會成因、以及其它校內英語學習相關活動

等。與會者對台灣的大學英語教育均留下了深刻的印象，並於會後提出許多問題，表達了深入瞭解的興趣。

當天在場聽眾含 TESL Director Prof. Robinson, ESL Director Prof. Madden, IEC Director Prof. Schwartz, 及另四位 TESL Program 專任教師，對台灣英語教育實施現況與美國的相似程度深感興趣。藉由這個機會，我認識了多位同行，成為他們對台灣英語教育瞭解的窗口，也開啟了未來進一步交流的可能。

2. 拜訪 St. Cloud State University TESL / ESL / IEC Program 及 St. John's / St. Benedict University

經由 Professor Robinson 的協助，此行有機會與同行多位教師有個別談話瞭解彼此工作內容的機會。其中，Robinson 為非常資深的 TESL Director，對整個英語教學人才培養及課務推動有數十年的資歷，加以其個人跨文化婚姻的經驗，對亞洲的韓國文化有很深的接觸，並因此衍伸出美韓間大學的正式跨校交流，包括教師與學生互訪等活動。Robinson 對台灣接觸不多，但十分有興趣，提出很多問題，以瞭解我們在台灣教英文的運作方式及台美間合作交流的可能性。我則對於他們的 TESL 母語學生是否有機會到台灣交換並擔任大學英文課程助教很感興趣，這將會對本地大學生在學英文方面帶來很好的刺激，惟政大外文中心並非學系，沒有隸屬自己的學生，故也沒有正式討論交換學生的機會，我則允諾將相關訊息帶回給本校英文系 TESL 相關教師。畢竟相隔時差 14 小時的兩地大學，能有交流的機會也不容易，我們都抱著開放的心態多瞭解彼此的工作，希望

能有機會促成進一步深入的學術交流。

ESL (English as a Second Language) Director – Professor John Madden 也與我討論並撥冗為我介紹 ESL Program 的運作，比較特別的是他們將 TESL 研究生的學習與 ESL 的教學連結在一起，雖然在行政管理上年年有新學生要帶，才能確保 ESL 的教學品質，但這樣的實習方式可讓教學合一，且在同一教學場域進行，頗有值得借鏡之處。

IEC (Intensive English Center) Director – Professor Michael Schwartz 曾在約二十年前在台灣中部教過英文，目前負責 IEC 的運作，與 ESL 不同的是，IEC 的學生程度一般較低，不少是近年來人數快速增加的大陸學生，若還不到進 ESL 的程度，就須先在 IEC 上課，逐級升到必須的程度後，才能進入 ESL，之後再進入正式大學學制。它的運作方式有些類似本校的華語文教學中心，課程密集，也搭配不少課室外實際運用所學的機會。在我們討論的過程中，有不少年輕學生來尋求各種課業上、生活上的幫助，看來在 IEC 輔導工作還是很重要的一環。

在 St. Cloud State University 附近還有一家私立教會大學，是專收男學生的 St. John's 及專收女學生的 St. Benedict，因在美研究伙伴對這所據稱年學費四萬美元的優質小型大學很推崇，故我造訪了其 Education Department 的 Professor and Chair – Dr. Janet Grochowski，做了簡短的會談與心得交換。Grochowski 也是一位十分資深的教授，他們在這所學校特別注重教學品質及學生輔導，讓我實際感受到追求 SSCI 論文之外的一種高等教育思維。

3. 參與 Graduate Assistant Meeting

這個 GA Meeting 其實是 TESL MA Program 的一部份，修課學生都必須在 ESL 上課，上課的同時也要在學期中定期共同討論教學，Professor Madden 邀請我參與一場 GA Meeting。當時正值新學期 Spring term 之始，有十五位左右的研究生成員，由 Prof. Madden 帶領，他指定學生鎖訂 ELT Journal，在這份期刊中依自己有興趣的主題找一系列文章來讀，並做課堂分享報告。我對於選定一份期刊的做法感到十分認同，且其選擇 ELT Journal 而非其它較為研究取向的期刊，可看出安排者的思維是聚焦在有研究基礎的教學實務，而非純研究。

ELT Journal 向來是我的重要讀物，個人的心得是 ELT Journal 文字十分淺白易懂，文章也多在五千字以內，每一篇的參考文獻多半只有十筆左右，以教學研究為出發點，但特別強調文章要對教學實務有所啟發或貢獻，不要冗長的研究程序報導。其實要發表在 ELT Journal 的文章必須要能將研究、理論融會貫通，才有可能以短篇幅深入淺出做討論。如果研究生的未來發展方向不在做研究而是進行高品質的教學，ELT Journal 實在是很好的選擇，學生經由老師的帶領，熟悉一份優質期刊的歷史資料，並在同時運用到自己進行教學實習的 ESL 課堂上，再加上同儕間的討論與腦力激盪，相信對於教學能力的培養有相當的幫助。

對照於自己過去在他校英語教學所遇到的碩士班學生抱怨，稱在研究所所學對未來在中學教書助益不大，這樣的安排與做法實可不失學術基礎，又能藉由廣大作者群給學生在教學理論與教學實務上很好的導引。

4. 課室觀察 SLA

參與 Prof. Madden 的 Second Language Acquisition 碩班課程，與過去在美國的相關課程經驗十分相似。

5. 對 St. Cloud State U 大學部學生介紹台灣大學生生活

經研究伙伴介紹，應當地學生社團 Geography Honors Society 之邀，利用週五中午時間做一場非正式的演講與分享，主要是當地大學生希望有機會多認識各國訪客及其文化。我依自己所熟悉的內容，擬定題目為 College Life in Taiwan，先從台灣大學生學英語的情形談起，順便自剛完成的學生期末考口說錄音檔取材，播放學生的英文口說內容，碰巧當時考試的題目有對美國 911 事件的瞭解與敘述、颱風及其對人們生活影響的介紹、對美國大學生活的想像等，都頗適合與美國大學生分享，也成為他們初步認識台灣大學生的途徑。此外，我準備了一些校園照片，包括校園機車停車場大批機車的壯觀景象、圖書館中的蔣公銅像、外文中心舉辦的國際學生文化交流工作坊現場、學生在行政大樓前舉牌抗議研究生助學金被刪等，最後以 Youtube 上政大學生的校園跑酷短片作結，演講大約三十至四十分鐘，隨後由在場學生自由發問。一如預期，學生問題很多很踴躍，其中也有幾位曾到過台灣或十分有興趣在未來到台灣看看，我留下了連絡方式予他們，未來他們若真有機會到訪，很可能創造英文課程中學生良好的學習機會。

二、 研究成果

此行中除延伸的訪問與拜會之外，最主要的是與在當地教書的朋友進行討論。個人計畫擬訂之初欠缺相關經驗，把跨國蒐集研究資料想得太簡單，

其實很多時候，視國情不同與研究倫理規範，要以研究計畫通過當地的審查或相關會議，方能進行問卷或訪談，此外被研究者的接受度也是一個重要的因素。這次未能如預期大量蒐集資料，主要是因研究伙伴的領域並非 TESL，其主要的接觸對象並不十分適合，但也透過對方的幫忙與安排，才有機會跟 TESL 相關的教研人員做了不少深入的討論與接觸，以第一次造訪當地而言，成果尚屬理想。藉由各自的介紹與相關教研活動的參與，不少當地教師都表示兩處的教學環境與教學實務有很多相似之處，也表達了進一步加強合作與交流的意願。在我回來至今近一個月，與 TESL Director Prof. Robinson 陸續有 email 往來連繫，相信繼續努力可以有更具體的研究成果。

過去近兩年計畫期間，幸有國科會的資助，得以至不同國度進行英語教學相關研究，也因此開拓不少眼界，建立不少關係，順道邀請了學者來台在政大外文中心主辦的研討會演講、開啟未來學術交流及進一步合作的可能，在高等教育越來越無國界的今日，個人在研究、教學及行政各方面都有豐富的收穫。

三、 建議

1. 本校對外宣傳的文宣品應可製作更為精美且全面，自美國進入政大網站英文版的內容也不足，都是可以改進的地方。此外，此次並發現英文授課課程是外國學生來台交換的重要考量。Prof. Robinson 有相當多的國際交流經驗，對多數中國大學不能提供他們交換生英語授課課程感到很失望，在看到政大文宣品中英文授課課程清單後，對未來的實質合作感到十分興奮。這也是目前政大及國內許多教育單位正在努力的方向。

2. 這次接觸到的多是有 PhD 學位且在教學型大學工作的教師，言談間感受到他們對教學的重視與驕傲，並認為這是相較於當地研究型大學的優勢，一點沒有研究不如人的焦慮，令我感觸良多。

3. 有教師提起其所開授的線上課程，所有的講授均已數位化，老師的工作重點在設計優良的課後作業、仔細研讀學生的作業並給予個別回饋。這真是一個未來避免不了的趨勢，標準化的流程都可由機器代勞，老師的心力更要投注到個別的關注與導引，這對老師是更大的挑戰，但也更符合我近年研究的 Assessment *FOR* Learning 的理念，惟有花時間瞭解學生所知與不足，才能給予適切有用的教導，要做到這樣，老師要更為知識淵博、經驗豐富。課程線上化、教學個人化或許是即將看到重要改變。

四、 其他

攜回資料有下

- 所觀察課室之課堂講義兩份
- St. Cloud State University TESL Program 海報兩張
- St. John's/St. Benedict 之 DVD 一片

附件五

國科會補助專題研究計畫出席國際學術會議心得報告

日期：102 年 7 月 16 日

計畫編號	NSC100－2410－H－004－186－MY2		
計畫名稱	學習動機、課室評量與教師權變		
出國人員 姓名	黃淑真	服務機構 及職稱	政治大學外文中心副教授
會議時間	102 年 7 月 8 日至 102 年 7 月 11 日	會議地點	希臘雅典
會議名稱	(中文) 雅典教育研究機構第六屆語言教學與語言學研討會 (英文) The 6th ATINER on Language and Linguistics (Athens Institute for Education and Research)		
發表題目	(中文) 以目標設定融入形成性評量寫作修改教學的成效檢驗 (英文) Effectiveness of Goal-setting in EFL Writing Revision		

一、參加會議經過

7/5 晚自台北出發

7/6 晚抵希臘雅典

7/7 休息準備

7/8-11 參與會議

7/12 啟程返國

7/13 抵台

二、與會心得

本次研討會經驗與過去略有不同，分為以下幾點說明：

1. 與會者國籍非常多元：過去個人多參加美國地區或亞洲如日本的大型研討會，此次或因希臘之地緣關係，有機會接觸一些不同國家的學者，除較常接觸的美、加、中、港、日、韓、德、澳、紐之外，與會學者來自地區如 Israel, Turkey, Croatia, Czech Republic, Ukraine, Belgium, Italy, Russia, Romania, Georgia, South Africa, Bahamas 等，聆聽這些地區學者在相關學術領域所做的研究報告是非常新鮮有趣的經驗。

2. 聽取的報告內容多元：過去參與的大型會議因場次分類較細，加以個人研究逐漸專注於固定領域之後，總會選擇自己熟悉的題目去聽，而這次會議的場次較有限，在會場中聽取了多個平常較不會接觸的題目，例如 language confluency, negotiation of meaning, grammatical morphemes, Chinese particles, universal grammar, indigenous languages 等，也都是過去在博士養成時期涉獵過的內容，有機會看看這些領域較新的研究如何進行也是很好的思考訓練。

3. 主持了一個場次共六篇論文：已多年沒有擔任主持人的任務，這次受派欣然接受，時間掌控十分良好。

三、發表論文全文或摘要

Athens Institute for Education and Research **Abstract Submitting Form**	
Conference	6th Annual International Conference on Languages & Linguistics , 8-11 July 2013, Athens, Greece
Title of Paper	Effectiveness of Goal-setting in EFL Writing Revision
For more than one author, please copy and paste the following eight rows for each additional author.	
Title	■ Dr □ Mr □ Ms □ Other Specify:
First Name	Shu-Chen
Family Name	Huang
Position	Associate Professor and Director
Affiliation	Foreign Language Center, National Chengchi University
Country	Taiwan
E-mail	huang91@nccu.edu.tw
Telephone(s)	886-2-29393091 ext. 62027/88126
Fax	886-2-29387363
Abstract	To facilitate learners' writing revision process, findings in goal-setting literature were sought and designed into an EFL writing course for college students. This study examined whether the provision of goals or goals plus a strategy list may be facilitative to EFL college students' performance in revising their own writing drafts. Three intact groups were assigned to one of the three conditions: control, goal-setting, and goal-setting + a Strategy List. All learners went through the same draft-instruction-revision sequence for three tasks that focused on fostering an awareness of performance criteria and self-assessment. The three groups were differentiated in the worksheets they used when revising their drafts. The first group was told to revise and only reflected on the strategies they used after the completion of revision. The second group was guided to set personal goals before revision. The third group was facilitated by both goals and a list of strategies. The completed drafts and revisions were graded by two independent outside raters. Effectiveness of the three types of goals on overall and specific writing quality was examined using ANOVA and MANOVA. Learners' experiences with these three goals were also probed in focus group interviews. Results indicated that significant improvement existed in all nine draft-revision pairs. The three groups did not differ in the scale of such improvement on each single revision. But when all revisions were considered together, the third group outperformed the first and the first outperformed the second. Such difference was not observed when only drafts were compared across groups. Implications of goal-setting in teaching revision were discussed. It is hoped that the field would be better informed on how to use goals in promoting assessment for learning with the help of practical goal-setting mechanism.
Keywords	English as a foreign language, writing revision, goal-setting

Please email to: atiner@atiner.gr as an attached file or fax it to +30 210 3634209

四、建議

無

五、攜回資料名稱及內容

會議議程、另有發表論文摘要集已由大會公布於網站

http://www.atiner.gr/languages.htm

六、其他

無

Effectiveness of Goal-setting in EFL Writing Revision

Shu-Chen Huang
National Chengchi University
Taipei, Taiwan

The 9th Annual International Conference on Language and Linguistics
8-11 July 2013 Athens, Greece

1

Outline

- The Concept of Assessment for Learning (AfL)
- Ways of putting AfL into practice in an EFL writing context
  - An AfL Instruction Base
  - Monitor with Goal-setting
  - Scaffold with Strategy Provision
- Results: A weak claim
  - AfL helps learners improve their drafts
  - AfL + Goal + Strategy > AfL > AfL + Goal
- Discussions and Implications

2

On Assessment

Summative Assessment
- High stakes
- Testing theories
- Reliability and Validity
- Bell curve, normal distribution

Formative Assessment

In the classroom, teacher-based
Assessment *of* Learning
Assessment *for* Learning

3

Assessment to Serve the Purpose of Learning and Teaching

- "Central to AFL is the integration of assessment into instruction as an ongoing process, where assessment information is used by teachers to adjust their teaching and by students to adjust their learning processes." (Popham, 2008)

4

Assessment FOR Learning (AfL)

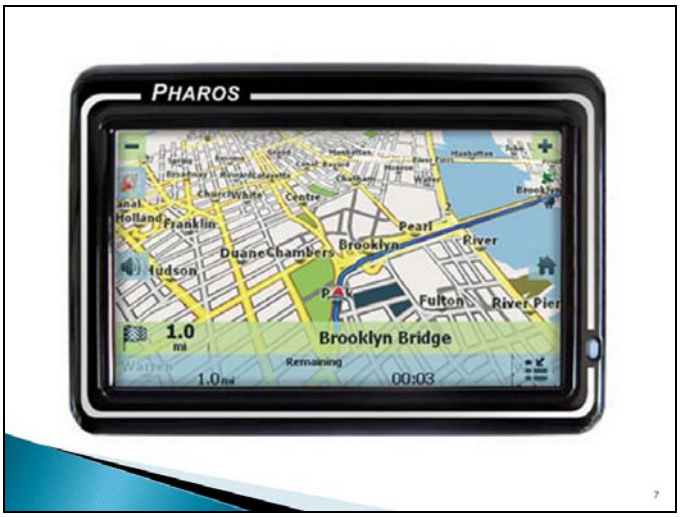
- AFL is characteristic of an assessment culture, that builds on scaffolding, aligning instruction in support of learning, and providing students with greater autonomy in their learning processes (Black & Wiliam, 1998a).

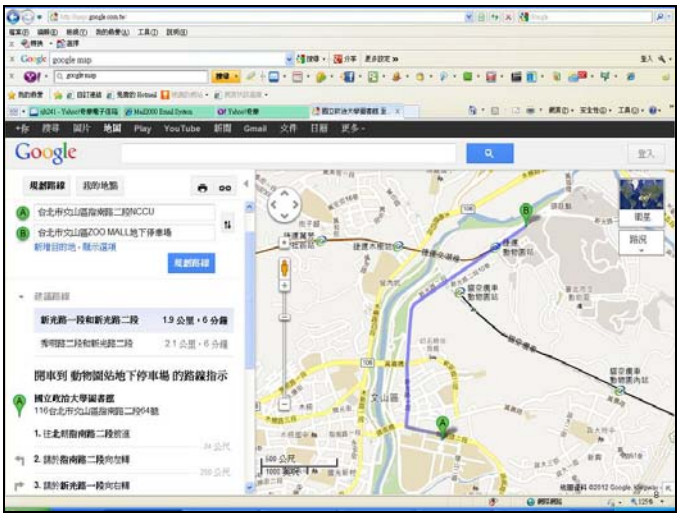
5

Black & Wiliam's Formative Assessment Model

	Where learner is going	Where learner is right now	How to get there
Teacher	1. Clarifying learning intentions and criteria for success	2. Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding	3. Providing feedback that moves learners forward
Peers	Understanding and sharing learning intentions and criteria for success	4. Activating students as instructional resources for one another	
Learner	Understanding learning intentions and criteria for success	5. Activating students as the owners of their own learning	

6





Putting AfL Ideal into Practice

- What can practicing teachers do?
- AfL keys are to "monitor & scaffold." (ARG, 2002; Pat-El, Tillema, Segers, & Vedder, 2013; Stiggins, 2005)

Put it into practice

- An AfL-based instruction alone
- + Monitor - Goal-setting
- + Scaffold - Strategy Provision

9

Goal-setting Theory

- The theory concerns human behavior and motivation as well as how efforts could be optimally directed under appropriate goals.
- This theory has been applied in areas that require goal-directed self-regulation to enhance effectiveness and efficiency (Locke & Latham, 1990).
- Areas applied: workplace management, athletes' sports training, patients' health management & therapies, and many other human behaviors

10

Goal-setting Mechanisms

- When a goal that is reachable and specific and to which people are committed, they will automatically utilize goal-relevant knowledge and skills to achieve such a goal (Latham & Kinne, 1974).
- When existing knowledge and skills are inadequate, people will search for similar strategies from their past experiences and consciously regulate their behaviors (Latham & Baldes, 1975)

11

Goal-setting Mechanisms

- 3 factors have been identified as influencing the effect of goal-setting
  - individual commitment to goals,
  - performance outcome & feedback information,
  - task complexity (Latham & Locke, 2006; Locke & Latham, 1990, 2002, 2006).
- Goal setting has rarely been used in academic learning situations, and the results to date have been mixed.

12

Caution

- When strategies and skills that may be used to achieve goals are not yet automatized, the presence of specific goals has the potential to cause anxiety and may be worse than the complete absence of goals.

13

L2 Learning Strategies

- L2 learning strategies are the learner's goal-directed actions for improving language proficiency or achievement, completing a task, or making learning more efficient, more effective, and easier. (Oxford, 2011)
- Growing interest in incorporating a focus on learning strategies into language curricula

14

Strategy Instruction/Assistance

- Learning strategy instruction were most successful when they were taught in context. (Hattie et al., 1996)
- Boekaerts, Pintrich & Zeidner's (2000) *Handbook of Self-regulation* emphasizes strategies such as setting goals.

15

Revision Strategy Instruction

- Butler & Britt (2011) taught revision strategies
- They found that revision, especially global revision, could be improved by explicit instruction.
- The researchers argued that such revision instruction was especially critical for learners who had been accustomed to having their writing corrected for grammar and local errors only.

16

Goal-setting in Academic Learning Situation

- Zimmerman & Kitsantas (1999) on revision
  - Outcome-goal group: focusing on minimizing the final number of words
  - Process-goal group: emphasizing a three-step method
  - Shifting-goal group: shifting sequentially from process to outcome goals
- Shifting-goal group outperformed the process-goal group, which in turn outperformed the outcome-goal group

17

An Instruction with AfL Principles

- Basis: Instruction abiding by AfL principles
- Integrate additional features to facilitate learner behavior – Is more better?
- The 3 groups
  - Control: AfL-based Instruction Only
  - Monitor: + Goal-setting
  - Monitor & Scaffold: + Strategy Provision

18

	1st session of 50 minutes		2nd session of 50 minutes	
Week 1	Individual writing of the draft for 30 minutes	Teacher-led discussion of a TA sample	Teacher modeling peer review using the TA sample and instructional rubrics	Peer review in small group. At the end, learners collect reviews from peers and self-assess
Week 2	Teacher's summary of assessment on drafts and explanations of problems indentified	Teacher-led discussion on how to revise; Group discussion on specific revisions	AfL Only Group	Individual revision for 30 minutes; Reflection on revision using worksheets
			AfL + Goal; AfL + Goal + Strategy	Planning revision using worksheets: AfL + Goal (set goals); AfL + Goal + Strategy (goals + a revision checklist); Individual revision for 30 minutes; Final check against work sheets

Figure 1. Instructional Steps of One Draft-Instruction-Revision Unit in the Three Groups

19

Research Questions

- Did student writing improve from each draft to its revision? Were there differences among the three groups?
- In view of all three draft-revision cycles, did student writing improve in terms of the quality of drafts and revisions? Were there differences among the three groups?

20

Appendix A: Revision Worksheet – AfL Only Group

Peer- and Self-Assessment

Items	Draft – peers (15----------1)	Draft – self (15----------1)	Revision – self (15---------1)
Holistic Score			
Argument			
Organization			
Lexical Use			
Grammar Structure			

Reflection of revision

These are the things I tried when I revised my draft	Done ✓

21

Appendix B: Revision Worksheet – AfL + Goal Setting Group

After completing the draft, peer review, and class discussion on revision, please select one or more items from the rubrics as your goal for revision.

Revision Goals: H, A, O, L, G (Please circle one or more goals)

After you have decided on your goals, list the strategies you have learned and could be used in your own revision before you start. Once you have finished your revision, check in the box to the right those strategies you have actually used.

Revision Strategies (List the things you will do in revision)	Checked ✓

Peer- and Self-Assessment

Items	Draft – peers (15----------1)	Draft – self (15----------1)	Revision – self (15---------1)

22

Appendix C: Revision Worksheet – AfL + Goal Setting + Strategy Provision Group

After completing the draft, peer review, and class discussion on revision, please select one or more items from the rubrics as your goal for revision.

Revision Goals: H, A, O, L, G (Please circle one or more goals)

After you have decided on your goals, list the strategies you have learned and could be used in your own revision before you start. Once you have finished your revision, check in the box to the right those strategies you have actually used. A strategy list is provided below for your reference.

Revision Strategies (List the things you will do in revision)	Checked ✓

A sample list of revision strategies for your reference

Items	Possible Strategies
1. Review and analyze the draft, find problems and areas in need of revision, in order to improve the quality of writing.	Read to yourself your draft from the beginning to the end, or have a peer read it to you. As you move along, underline places that sound strange or could be improved.
2. Check your arguments, thesis and topic	If a problem is identified, ask yourself what you really want to

23

Descriptive Statistics of All Drafts and Revisions by All Three Groups

		AfL Only (n=40)			AfL + Goal (n=35)			AfL + Goal + Strategy (n=33)		
		M	*SD*	95% CI	*M*	*SD*	95% CI	*M*	*SD*	95% CI
1st	Draft	8.16	1.74	[7.60, 8.72]	8.67	1.69	[8.09, 9.25]	6.73	2.29	[5.91, 7.56]
	Revision	9.36	1.76	[8.78, 9.95]	9.66	1.79	[9.01, 10.32]	8.54	2.09	[7.73, 9.34]
2nd	Draft	9.50	1.29	[9.07, 9.93]	8.73	1.44	[8.22, 9.24]	7.61	2.30	[6.78, 8.44]
	Revision	10.23	1.30	[9.78, 10.68]	9.60	1.08	[9.15, 10.05]	8.98	2.17	[8.19, 9.78]
3rd	Draft	7.81	2.25	[7.03, 8.59]	6.65	2.56	[5.71, 7.58]	8.43	1.43	[7.89, 8.98]
	Revision	8.28	2.01	[7.63, 8.93]	7.89	2.43	[6.84, 8.94]	9.56	1.34	[9.02, 10.10]

24

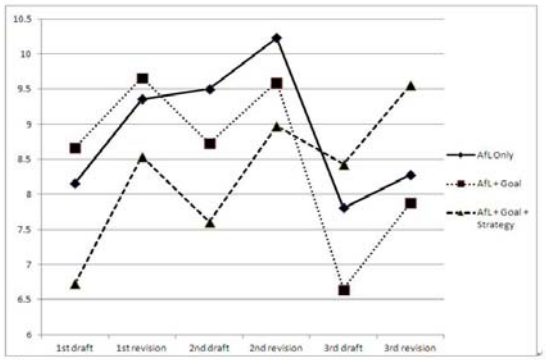


Figure 2. Line Graph of the Writing Scores for Three Tasks for the Three Groups

25

Descriptive Statistics of Draft-to-Revision Improvement of 3 Groups on 3 Tasks

	1st Writing			2nd Writing			3rd Writing		
Groups	*M*	*SD*	95% CI	*M*	*SD*	95% CI	*M*	*SD*	95% CI
AfL Only	1.22	0.83	[0.94, 1.49]	0.68	0.7	[0.43, 0.92]	0.62	0.71	[0.37, 0.86]
AfL + Goal	1.18	1.14	[0.76, 1.60]	1	0.82	[0.66, 1.34]	0.8	1.19	[0.29, 1.32]
AfL + Goal + Strategy	1.71	1.65	[1.08, 2.35]	1.37	1.83	[0.70, 2.04]	1.04	0.87	[0.69, 1.39]

26

Comparisons among 3 Groups

- On 3 Draft-Revision Improvements
  - An immediate effect
- On 3 Drafts
  - This could indicate that the learners had internalized what they had learned in the previous round and transferred this learning in new assignments.
- On 3 Revisions
  - This could indicate that learning to revise and accumulating this learning experience (under three treatments) over time caused the three groups to differ.

27

Table 4. Summary of ANCOVA on Draft-Revision Improvement among Three Groups in Three Tasks

ANCOVA Summary on the 1st Writing

Source of Variance	*df*	*F*	*p*	η^2	*obs. power*
Score of First Draft	1	14.34***	0.000	0.135	0.963
Group	2	0.22	0.803	0.005	0.084
Within Group Errors	92	(1.29)			

ANCOVA Summary on the 2nd Writing

Source of Variance	*df*	*F*	*p*	η^2	*obs. power*
Score of First Draft	1	0.51	0.479	0.006	0.108
Group	2	1.62	0.203	0.037	0.334
Within Group Errors	85	(0.94)			

ANCOVA Summary on the 3rd Writing

Source of Variance	*df*	*F*	*p*	η^2	*obs. power*
Score of First Draft	1	3.79	0.055	0.046	0.485
Group	2	0.84	0.437	0.021	0.189
Within Group Errors	79	(0.81)			

28

Table 5. ANOVA Summary on Three Drafts

Source of Variance	*df*	*F*	*p*	η^2	*obs. power*
Group	2	2.64	.077	.057	.512
Within Group Errors	87	(6.90)			

29

Table 8. ANOVA Summary of Three Groups on Revisions

Source of Variance	*df*	*F*	*p*	η^2	*obs. power*
Group	2	155.65***	.000	.812	1.000
Within Group Errors	72	(1.07)			

30

Table 9. Multiple Comparison on Revisions by Groups Using Bonferroni Tests

(I)Group	(J)Group	Mean Difference (I-J)	p	Difference 95% CI	
AfL	AfL + Goal	1.61***	0.000	1.18	2.04
	AfL + Goal + Strategy	-1.69***	0.000	-2.08	-1.29
AfL + Goal	AfL	-1.61***	0.000	-2.04	-1.18
	AfL + Goal + Strategy	-3.30***	0.000	-3.76	-2.84
AfL + Goal + Strategy	AfL	1.69***	0.000	1.29	2.08
	AfL + Goal	3.30***	0.000	2.84	3.76

31

Results

- On 3 Draft-revision improvements
  - No significant difference
- On 3 Drafts
  - No significant difference
- On 3 Revisions
  - Significant differences observed
  - AfL + Goal + Strategy > AfL Only > AfL + Goal

32

Discussions and Implications

- All revisions were better than drafts.
- Consistent draft-revision improvements
- This translation of AfL principles worked.
- It helped to explicitly teach learners how to improve their drafts.
- Similar to the findings of Sengupta (2000) and Butler and Britt (2011), the explicit instruction of revision is worthy of teachers' efforts.

33

Discussions and Implications

- The only significant difference found when
  - the revised versions were compared across groups
- AfL + Goal + Strategy > AfL > AfL + Goal
- When strategies and skills that may be used to achieve goals are not yet automatized, the presence of specific goals has the potential to cause anxiety and may be worse than the complete absence of goals.

34

Thank you!

huang91@nccu.edu.tw

35

國科會補助計畫衍生研發成果推廣資料表

日期:2013/10/03

國科會補助計畫	計畫名稱: 學習動機、課室評量與教師權變
	計畫主持人: 黃淑真
	計畫編號: 100-2410-H-004-186-MY2　　學門領域: 英語教學研究

無研發成果推廣資料

100 年度專題研究計畫研究成果彙整表

計畫主持人：黃淑真	計畫編號：100-2410-H-004-186-MY2
計畫名稱：學習動機、課室評量與教師權變	

成果項目			量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
			實際已達成數（被接受或已發表）	預期總達成數(含實際已達成數)	本計畫實際貢獻百分比		
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	1	1	100%		
		研討會論文	1	1	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	2	2	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	1	3	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	2	2	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)	無

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

■達成目標

□未達成目標（請說明，以 100 字為限）

□實驗失敗

□因故實驗中斷

□其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文：■已發表 □未發表之文稿 □撰寫中 □無

專利：□已獲得 □申請中 ■無

技轉：□已技轉 □洽談中 ■無

其他：(以 100 字為限)

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）(以 500 字為限)

本計畫至今共完成兩篇國際研討會論文發表，一篇 SSCI 期刊論文發表，一篇學術期刊論文投稿審查中，另仍有調查與訪談資料整理中. 研究成果提出課室評量的在地觀點，與國際最新研究發展對話，具英語教學領域之學術與應用價值.