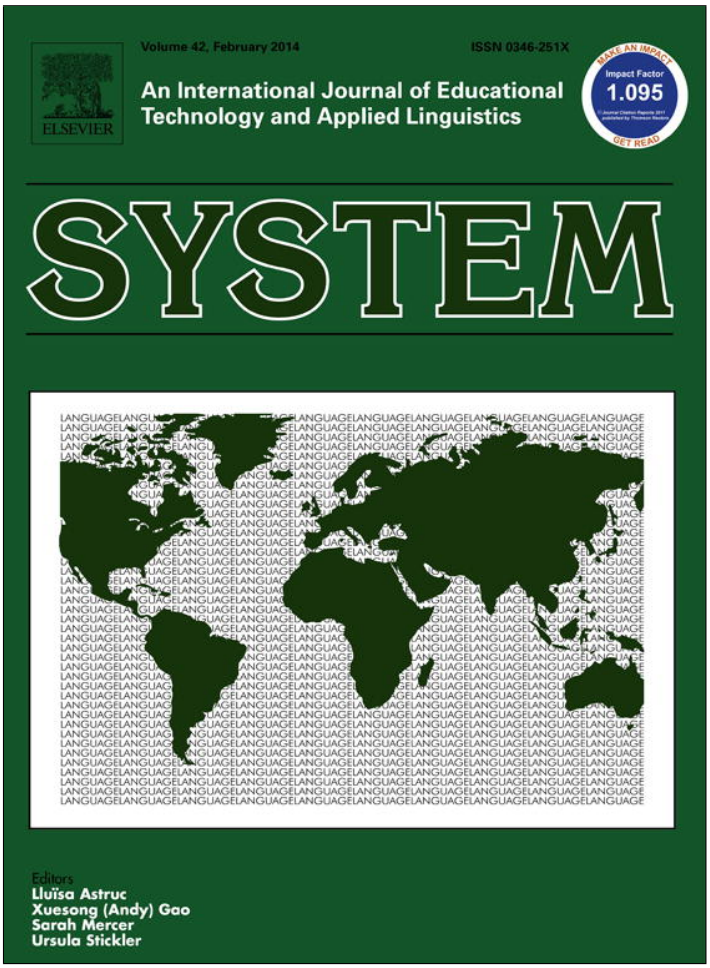

System 42 (2014) 333–343

Contents lists available at ScienceDirect

System

journal homepage: www.elsevier.com/locate/system

Student preferences vis-à-vis teacher feedback in university EFL writing classes in Japan

James A. Elwood [a,*], Jeroen Bode [b]

[a] Meiji University, Japan
[b] Tsukuba University, Japan

ARTICLE INFO

Article history:
Received 24 February 2012
Received in revised form 23 December 2013
Accepted 30 December 2013

Keywords:
Student preferences
Writing feedback
Feedback color
Feedback utilization
Writing revision
Affective reactions to feedback

ABSTRACT

While learning outcomes associated with type and frequency of feedback have been extensively researched, student perceptions of teacher feedback have received less attention. This manuscript reports on an investigation of student perceptions and preferences concerning teacher feedback in university EFL writing classes in Japan. Students generally reacted positively to feedback and exhibited strong preferences for detailed, handwritten feedback that addressed both content and mechanical errors. The color of feedback appears to be an issue of minimal concern as students indicated nearly equal preference for red and blue marking. Higher proficiency corresponded with lower anxiety levels, an increased willingness to ask questions about feedback, and more positive reactions to feedback, while the opposite was true for lower-proficiency students. Females preferred detailed, direct feedback more than males did, while males indicated somewhat higher anxiety concerning feedback. Finally, feedback was utilized only to a modest extent, a result that highlights the need to encourage and train students in its use. The study concludes with suggestions for further research.

1. Introduction

The current study began several years ago upon hearing an offhand remark from a veteran teacher: “My students don’t want feedback on their writing, and even if I give it to them, they don’t read it. Why bother?” Upon being asked how he knew those things to be true, that veteran teacher responded with the time-tested and unassailable answer that, as a veteran teacher, he “just knew”.

Being experienced writing teachers, we believed – or perhaps *wanted* to believe – that he was mistaken. Nonetheless, we felt compelled to investigate whether students wanted feedback, what types were preferred, and what was done with it. In studies conducted some 25 years ago in ESL contexts, Radecki and Swales (1988) and Leki (1991) found that students expected feedback and were disappointed when it was not forthcoming. Some 10 years hence, Ferris and Roberts (2001) and, in an EFL context, Lee (2004) obtained similar results. With the pace of change in the Internet Age, however, we wondered whether the subsequent interval might have witnessed changes in perceptions. With information readily available in cyberspace and editing easily accomplished with keystrokes instead of erasers, perhaps students’ feedback preferences and perceptions of feedback had come to diverge from those found in the earlier studies.

Learner beliefs play a crucial role in education. Inasmuch as teaching activities need to be perceived by learners as conducive to learning, foreign language educators should remain cognizant of learner beliefs or perceptions when planning

* Corresponding author. 1730-25 Hiregasaki, Nagareyama, Chiba 270-0161, Japan. Tel.: +81 4 7186 6434.
E-mail address: elwood.jim@gmail.com (J.A. Elwood).

0346-251X/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.system.2013.12.023

and conducting classroom activities. Learner beliefs have been identified as an important individual difference variable in L2 learning in general (Dörnyei, 2005; Kalaja & Ferriera Barcelos, 2003). In specific areas, L2 learner beliefs have been found to correlate with strategy use (Sakui & Gaies, 1999; Yang, 1999), motivation (Oxford, 1994; Pintrich, Marx, & Boyle, 1993; for an overview, see Horwitz, 1999), proficiency (Mori, 1999; Peacock, 1999; Yang, 1999), learner anxiety (Levine, 2003), and autonomous learning (Cotterall, 1995; Kalaja & Ferriera Barcelos, 2003). Moreover, teachers' classroom activities can be influenced by learner beliefs (Borg, 2003; Burgess & Etherington, 2002), with, for example, unrealistic beliefs or misconceptions about language learning adversely affecting the learning process (Sawir, 2002). As Schulz (2001) succinctly noted:

> Given that discrepancies in student and teacher belief systems can be detrimental to learning, it is important that teachers explore their students' perceptions regarding those factors believed to enhance the learning of a new language and make efforts to deal with potential conflicts between student beliefs and instructional practices. (p. 244)

Because a vigilant awareness of students' perceptions is crucial to successful learning outcomes, students' perceptions of feedback on their writing should be addressed. The following sections thus explain the various aspects of feedback addressed in this study.

1.1. Literature review

1.1.1. Feedback definitions

For the purposes of the current study, the various types of feedback queried are defined as follows. Direct (corrective) feedback is defined as "the provision of the correct linguistic form or structure above or near the linguistic error," which could include "the crossing out of an unnecessary word/phrase/morpheme, the insertion of a missing word/phrase/morpheme, or the provision of the correct form or structure" (Bitchener, 2008, p. 105). More succinctly, direct feedback means that the teacher explicitly provides the correct form.

On the other hand, indirect (corrective) feedback notifies the writer that an error has been made without explicitly correcting the error. As researchers of oral L2 production have found, learners must first 'notice' (Schmidt, 1990) that an error has been made. The same is likely true for written L2 production, inasmuch as failure to notice an error would logically preclude addressing it. The noticing process is also more complex than a simple dyad (i.e., notice vs. fail to notice) suggests, for the quality of noticing also appears to play a role in how the noticed discrepancy is addressed (Qi & Lapkin, 2001).

This feedback genre includes several methods, one of which is noting in the margin the number of errors in a given line. This could be of limited use, however, for errors can be of various types simultaneously, and a certain level of proficiency would be necessary to find the numbered errors. A second incarnation is noting the location of the error by underlining or circling it (Ferris & Roberts, 2001; Robb, Ross, & Shortreed, 1986). Although these methods indicate the presence or the location of an error, neither method shows the type of error. This can be remedied by the use of a numeric or orthographic code to indicate the type of error and thus provide further information (e.g., Kubota, 2001). This particular method thus lessens the cognitive load on the student and quite likely also reduces the time and effort expended, yet it requires the student to provide the correct form. The crux of these techniques is that students must identify and then correct the problem from the information provided and their own knowledge.

1.2. Various facets of feedback

The choice of a particular method of providing feedback is of importance, and the mode of providing feedback plays an important role, too. In the following section we will explicate some of the various components inherent in feedback on written EFL and ESL student work. Among these components are the degrees of explicitness (direct vs. indirect), timing (immediate vs. delayed), the manner of delivery (e.g., handwritten vs. delivered using technology), the source (self, the teacher, or peers), and even the visual presentation (i.e., the color of feedback). Inasmuch as the scope of the current study includes the degree of explicitness, the manner of delivery, and the visual presentation, in the following sections we will address these three factors.

1.2.1. Degree of explicitness

The relative merits of these modes of providing feedback have been discussed at some length in the literature on second language acquisition (SLA). Direct feedback, for example, seems to be preferred by teachers and students (Ferris & Roberts, 2001; Komura, 1999; Timson, Grow, & Matsuoka, 1999); although offering direct, detailed comments on individual papers is a time-intensive task for the teacher (Chandler, 2003), it reduces both the time and cognitive load required of students, whose primary task then becomes "reading the teacher's handwriting" (Duppenthaler, 2002, p. 132). Expeditious use of time is an important consideration in university EFL teaching in Japan, in which students typically have any given language class just once a week for 90 min.[1] Furthermore, direct feedback reduces other potential areas of misunderstanding, namely, confusion about the location or type of the error or about the meaning of error codes used. As Ferris and Roberts (2001) noted, such confusion can easily occur with lower-proficiency learners. Moreover, students sometimes feel that indirect feedback does not provide them with sufficient information to resolve more complex errors such as idiosyncratic and syntactic errors (Leki, 1991).

[1] Although nearly all universities in Japan utilize 90-min class periods, at the venue in this study classes were 75 min.

Although these arguments favor the use of direct feedback, providing indirect feedback rather than direct feedback also has merits. Lalande (1982) and James (1998) explained that indirect feedback requires learners to engage in guided learning and problem solving, processes that facilitate the type of reflection likely to foster long-term acquisition. Providing further indirect feedback has the potential to push learners to engage in hypothesis testing, a process that can induce deeper internal processing and promote the internalization of correct forms and structures (Doughty & Williams, 1998; Ferris, 2002; Higgins, Hartley, & Skelton, 2002; Vickerman, 2009).

Several caveats deserve mention, however. First, the positive effects of expending greater cognitive effort in using indirect feedback to make corrections might be offset by the cognitive load being so demanding that it adversely affects function by consuming an inordinately large proportion of working memory (Kellogg, 1996, 1999; McCutchen, 1996). Frustration can accrue when a student knows an error has occurred yet is unable to correct it in an expeditious manner (Leki & Carson, 1994), and delays can occur in knowing whether hypothesized corrections are correct (Chandler, 2003). In lieu of focusing on the various pros and cons of direct and indirect feedback, in the current study we investigated our students' perceptions about various facets of direct and indirect feedback.

1.2.2. Manner of delivery: handwritten vs. e-feedback

In addition, in a nod to the continuing advance of technology, students were asked whether they preferred handwritten or electronic feedback (hereafter, e-feedback[2]), and whether they preferred feedback written on their document or on a separate piece of paper. While the manner of delivery is at first glance a minor point, the use of computers and the concomitant provision of e-feedback have become commonplace. Among the studies that have investigated the efficacy of e-feedback and its numerous advantages, Liu and Sadler (2003) examined the effect of peer review in electronic and traditional modes on L2 writing. Results indicated that e-feedback corresponded with a larger overall number of comments and a higher percentage of revision-oriented comments, which resulted in more revisions by the technology-enhanced group than by the traditional group. However, the authors cautioned that the nonverbal aspect of face-to-face communication is indispensable in intercultural communication in a peer review setting.

On the other hand, Brodie and Loch (2009) found that handwritten comments were preferred to e-feedback by both students and evaluators. The handwritten feedback was perceived as more personalized and specific and therefore of greater benefit. However, the students indicated having taken very little action with the feedback provided.

In another study that addressed the use of e-feedback vs. traditional handwritten feedback, Denton, Madden, Roberts, and Rowe (2008) found that e-feedback was superior for several reasons. These included better transparency of the marking scheme, improved feedback legibility, and the inclusion of information on deficient points. In addition, e-feedback was found to require less time, which allowed the manuscripts with feedback to be returned to the students in a more timely fashion, an important consideration because "the speed with which appropriate feedback is provided to students is essential to learning" (Jones, 1996, ¶13).

As noted above, handwriting prowess can be a factor, and the use of e-feedback obviates the need to decipher teacher handwriting. On the other hand, e-feedback can lack the human touch that accompanies handwritten feedback on a document; as Ware and Warschauer (2006) noted, the importance of social interaction in feedback is an important consideration.

1.2.3. Visual presentation: color of feedback

Another aspect of the mode of feedback is the actual physical presentation of the feedback, which has received little attention in the SLA literature and represents an area of possible concern as a mediating factor in the uptake and effectiveness of feedback. One visual element is the color of the feedback. The lead author has for the duration of his teaching career used blue pens when providing handwritten feedback, partly as a matter of personal preference and partly to avoid the use of red,[3] which has been shown to trigger an avoidance reaction (Mehta & Zhu, 2009) and over time become associated with failure (Elliot, Maier, Moller, Friedman, & Meinhardt, 2007). The use of red has been shown to adversely affect cognitive performance, whereas the use of other colors (specifically, black, white, gray, and green) in a North American ESL context did not. On the other hand, blue appears to engender an approach motivation, which should facilitate involvement with the problem at hand (Elliot et al., 2007).

Contradictory results exist, however, in which red was found to enhance cognitive task performance. Investigating the effects of different colors in an office environment, Kwallek and Lewis (1990) found that participants made the fewest errors in a red office and the largest number of errors in a white office. In a simulated telemarketing task, Stone (2003) found that, over time, performance declined more in a blue venue than in a red venue, which suggested that "red is a stimulating color" (p. 63). This apparent contradiction might have been explained by Mehta and Zhu (2009), who found that red enhanced cognitive performance on detail-oriented tasks while blue enhanced performance on creative tasks. Interestingly, while their

[2] In this study, the two modes of delivering feedback are assumed to be essentially equivalent (i.e., the content delivered is equivalent), but just as an assumption of equivalence for different modes of evaluation should be treated with caution (see Meyerson & Tryon, 2003 for an in-depth discussion of equivalence criteria and studies), perhaps assumptions about feedback provided via different modes should also be handled cautiously (Ware & Warschauer, 2006).

[3] The authors have found this of particular concern with papers that have a considerable number of errors. When such papers are corrected with a red pen, the result can be a paper 'splashed' in red ink; based on our own experiences learning additional languages, we attest that such a flood of red ink can be quite discouraging for the learner!.

study was conducted in Canada, it included both North American and Chinese informants, indicating that their results could be valid for students raised in contexts outside of North America.

A related finding was reported in a study by Soldat, Sinclair, and Mark (1997) that dealt with the effect of paper color on mood and processing strategy. Among other results, they found that the extent of motivation predicts cognitive performance: In low-motivation contexts, participants doing tasks printed on blue and white paper outperformed the red-paper participants, while in the high-motivation condition there were no discernible effects. This result hinted that an environmental cue such as color can directly affect processing strategy in low-motivation participants. Second, for complex tasks, blue paper correlated with better performance than did red paper, while for simple tasks no difference was apparent. Interestingly, the red paper was perceived as conveying more positive affect then either the white or blue paper.

In Japan, however, the prevalent form of delivery for error feedback is a red pen; indeed, in Japan teachers and students alike generally carry red pens expressly for that purpose. This raises the issue of whether the use of red for error correction might be detrimental, as was found in North American and European contexts (Maier, Elliot, & Lichtenfeld, 2008). A repeated pairing of red with 'failure' might over time create a learned association between failure and the color red that diminishes any positive results from the feedback, and the use of red for correction also appeared to be linked with negative affective reactions that compromise learning. The use of more neutral colors such as blue or green, however, was perceived to positively influence affective reactions (i.e., to reduce anxiety).

However, generalizing to other contexts should be done with care, for colors might have different associations depending on the venue (Aslam, 2006). In Japanese contexts, for example, the use of red probably lacks such negation associations. As noted, red is the usual color used for error correction in schools. When personal seals (stamps) are affixed to official documents, using ink of a reddish-orange color is the accepted practice.[4] Furthermore, red in Japanese culture represents celebration, appearing with white on such items as wedding announcements and congratulatory communiqués, a pairing that occurs in some other Asian contexts as well (Madden, Hewett, & Roth, 2000). In spite of trends, however, considerable variation exists across different cultures (e.g., Silver et al., 1988; Wiegersma & van der Elst, 1988), so the perception of individual colors and of combinations thereof should be considered dependent upon context. Thus, the color preference for feedback was queried.

1.3. Additional considerations

1.3.1. Actions taken with feedback

Even if feedback is available and sufficiently transparent to be useful, the reality is that something must be done with it for it to benefit their writing. Such actions can vary from simply reading feedback (or not; see Brodie & Loch, 2009) to conscientiously incorporating it into revisions, perhaps as a result of consciousness-raising activities (e.g., Dheram, 1995). Whereas the effectiveness of revising based on feedback is generally considered a beneficial task, contextual factors such as programmatic and institutional attitudes toward writing and the classroom environment can mediate that effectiveness (Conrad & Goldstein, 1999; Goldstein, 2004). However, in this study we will content ourselves with the notion that feedback has a positive influence on revision and leave the issues of causality and learning outcomes to future research.

This again harkens back to the notion of deepening the learning experience (Higgins et al., 2002; Vickerman, 2009). In the current study one group of students had to revise their work, while the second group did not. As such, it was expected that the two groups would differ in their responses concerning actions taken after receiving feedback and perhaps also in their perceptions of feedback. The group for whom revision was required was expected to take more action with feedback, yet their perception of feedback was unclear: Would they perceive it as tiresome, or would it be considered an important, even valuable part of the class? On the other hand, the 'optional feedback group' was expected to utilize feedback to a lesser extent and perhaps view it more casually, namely, as the optional task that it represented for them.

1.3.2. Correlations with gender, proficiency, and anxiety

The final strand investigated in the current study was whether gender, proficiency, and anxiety played any roles in the perception of feedback. Although Zhang (1995) found no gender differences regarding feedback preferences with university EFL writers (nearly all from East or Southeast Asia and the Pacific islands), we were curious if gender differences might emerge for this particular sample. Second, inasmuch as color differences were found to correlate with affective reactions (i.e., red with negative affective reactions and blue or green with positive affective reactions such as lower anxiety; Maier et al., 2008), anxiety associated with the series of tasks was queried. Finally, we posited that writing proficiency would also play a role, with higher proficiency writers, for example, both more able and more likely to utilize feedback.

1.4. Feedback research in Japan

As briefly outlined to this point, feedback is a complicated, multi-faceted undertaking, and it has received some scrutiny in Japan, where the current study was conducted. Among the studies that have investigated perceptions of feedback and affective results of feedback, Rinnert and Kobayashi (2001) examined perceptions of feedback and found that low-level

[4] However, red is not used exclusively. Blue ink is used to certify or attest a document, as when a translator affixes his or her seal to a translation of an original document.

students were concerned with content whereas higher-level students scrutinized organization and textual devices. In a very extensive study conducted with Japanese university students, Timson et al. (1999) found that a large majority felt error correction to be both necessary and wanted to facilitate gains in fluency and that such correction should be direct. Moreover, correction by instructors was preferred to that by peers. Duppenthaler (2002) reported on the effect of three types of written feedback on student motivation in a high school context in Japan. The primary finding reconfirmed the positive effect that journal writing had on motivation regardless of the type of feedback.

Researchers have also delved into empirical results of feedback. Looking at peer feedback in EFL writing contexts, Kamimura (2006) found positive effects on essay quality and length, but for fluency no significant improvement was found. Higher-level students tended to provide feedback on discourse-level concerns, whereas lower-level students addressed sentence-level issues. Several studies have examined various types of feedback in Japanese contexts, including Robb et al. (1986), who perused various types of feedback and found evidence against the use of direct feedback. However, in a study that addressed the effect of corrective feedback on the accuracy of English article use by Japanese university students, Ellis, Sheen, Murakami, and Takashima (2008) found that such feedback facilitated improvement of article use. Tono and Kanatani (1995) established that direct feedback on high school students' English writing benefitted high-proficiency students more than those of lower-proficiency levels, while Saito (2008) concluded that peer assessment is a robust system in which rater training enhances the depth of feedback, specifically in the number of comments and increased rater consistency. Saito and Fujita (2009) found "overall similarity" between peer and rater assessments although group cooperation did not strongly correlate with quality of group presentations, a result that should engender some caution when employing peer assessment.

Finally, feedback in hybrid contexts has also been the focus of research. A recent study by Bower and Kawaguchi (2011) investigated two types of feedback associated with a Japan–Australia exchange project in which students conversed using computer-mediated communication. Their findings included the effectiveness of both negotiated meaning during the computer-mediated communication (CMC) as well as explicit feedback provided to the students afterward.

2. Research questions

Based on the need to further investigate student preferences regarding feedback on writing and add to the literature on feedback in Japanese contexts, the following research questions underpinned the current study:

1. What mode of feedback do EFL students prefer?
2. What affective reactions do EFL students have toward writing feedback?
3. What color do EFL writing students prefer for feedback?
4. To what extent do feedback preferences vary by gender, anxiety level, and proficiency level?

3. Method

3.1. Venue

The data for this study were collected from extant writing classes at a national university in eastern Japan. First-year students at that university were required to take a year-long course on academic writing that met once a week for 75 min over the three 10-week terms for a total of 30 classes. The classes were conducted in English by the teachers, one a native speaker of English and the second a bilingual Dutch–English speaker. Both were veteran writing teachers with some 10 years of experience teaching academic writing to EFL learners. Over the course of the school year each participant had to write a series of academic essays, all of which were graded by the respective instructor and returned to the student with direct feedback.

3.2. Participants

Participants included 410 first-year students with an average age of 19.16 (SD = 0.65) in three cohorts over a three-year period at a national university in eastern Japan. During that period, all first-year students were required to take a year-long academic writing course that met once a week for 75 min. In Japanese secondary education, students are required to complete six years of English education, meaning the students in the current study were completing their seventh year of formal English education. These students represented eight majors and the top level of overall English proficiency at this university.[5] Females accounted for 44.63% ($n = 183$) of the participants, males accounted for 54.88% ($n = 225$), and two participants did not provide gender information.

[5] Incoming students were streamed based on results of an internally-developed placement exam; only students in the topmost A level took writing courses. Upon completion of their first-year courses, students were required to take a year-end exit exam that was administered within two weeks of when data were collected for this study; those test results comprised the objective measure of students' overall English proficiency. The psychometric performance of the exit exam has been examined carefully, with results exhibiting satisfactory reliability (Cronbach $\alpha = .813$) and a reasonable degree of concurrent validity (a correlation of 0.806 with the TOEFL-ITP). In very approximate terms, exam scores corresponded to TOEFL-IPT scores with a range of 524–609 and a mean of 570 (SD = 27), which suggests that most students were at the CFER B2 Level.

Of the 410 participants in this study, 49.3% ($n = 202$) were in classes in which students were permitted to hand in hard copies (i.e., paper copies) or send computer files via e-mail. Feedback was handwritten in blue ink on the paper submissions and done on a computer using the 'Track changes' function for the e-submissions; in 'Track changes' the default mode colors were retained, meaning deletions were noted in red and insertions in blue. In the first instructor's classes, students created nine original essays, for which revision was optional. The optional nature of revision was expected to appear as a significant difference in the actions taken after receiving feedback. Students of the second instructor were required to submit four academic essays as computer files. Direct feedback was provided using the default mode of 'Track changes' and essay revision was required, meaning that each of four unique essays was revised and resubmitted (for a total of eight submissions).

Although the two teachers utilized slightly different syllabi, the groups were tentatively treated as belonging to one population of first-year EFL writing students at this particular university. The two groups were then examined to determine whether they indeed formed a single population or were distinct populations.

3.3. Data collection and analysis

Data were collected at the end of the final class of the school year using a quantitative questionnaire developed for this study (see Appendix A) and an open-ended followup questionnaire administered four months later. The year-end questionnaire included 47 5-point Likert-scale items, on which a response of 5 indicated a high degree of the entity in question (e.g., "very competent") and a response of 1 indicated a low degree (e.g., "minimally competent"). The instrument was comprised of several sets of questions, the first four of which queried experience doing a particular task, perceived competence when doing so, perceived anxiety when doing so, and preference for certain types of written feedback. The next set asked about the type of feedback desired; the students were shown images of a paragraph with the respective types of feedback (i.e., direct feedback, various types of indirect feedback, and finally no feedback). The following sets of items investigated actions taken with feedback (e.g., revising the paper), the participants' affective reaction to feedback, and preferences for mode of delivery (e.g., preference for red ink or another color). The survey was quite short, with most participants completing it in about 10 min. The followup questionnaire consisted of seven open-ended questions soliciting opinions and comments about the various sections on the year-end questionnaire.

After the initial data screening, descriptive statistics were examined and the various items compared using *t*-tests. To minimize the possibility of Type-II errors, an FDR correction (Benjamini & Hochberg, 1995) was applied to each of the sets of questions. Rasch analysis (Rasch, 1960) was utilized to obtain interval-level person ability estimates for self-perceived writing proficiency and anxiety. The strata statistics derived from the Rasch person separation values[6] indicated that three proficiency groups and two anxiety groups could be formed; to reduce uncertainty about scores near the cut point while retaining a reasonable sample size, scores within one standard error of measurement of the respective means for proficiency and anxiety were removed, yielding groups of high ($n = 82$) and low proficiency ($n = 104$) as well as high ($n = 121$) and low anxiety ($n = 95$).

4. Results

The data from the three cohorts were screened for missing data, non-normality, and the presence of outliers. Of the original 433 respondents, 23 returned incomplete surveys that were excluded from the analyses. The results for the groups of the respective instructors were then examined to ascertain whether they belonged to a single population. In spite of having studied using different syllabi, the two groups exhibited few statistically significant results that reflected those syllabi differences. For example, in one syllabus, writing e-mail and longer (e.g., 5-page) essays played more prominent roles, which were reflected in the results: perceived competence, extent of experience, and anxiety differed by a statistically significant margin for both writing e-mail to the teacher and writing a 5-page report. The other difference of note was in the preference for red ink when receiving feedback: In the class in which feedback was provided in blue, students were predictably more favorably inclined toward its use. However, for all other tasks and queries – most notably the actions taken after receiving feedback – the two groups were statistically indistinguishable and thus treated as a single population.

The initial queries investigated the extent to which the participants had experienced a series of tasks, the extent to which they perceived themselves to be competent, and the degree to which they felt anxious. Predictably, the results indicated that students felt less proficiency and more anxiety with tasks in which they had less experience or as the tasks became more difficult and complex. For example, for writing a self-introduction, students reported the highest levels of experience (mean = 2.99) and competence (3.26) in addition to the second lowest level of anxiety (2.70). On the other hand, writing a 5-page report and writing a speech both were associated with lower rates of experience and competence and higher levels of anxiety (Table 1).

The following section addressed the type of feedback that participants desired. Detailed feedback was desired on both mechanical issues and content, echoing findings elsewhere (Fathman & Whalley, 1990; Leki, 2006; Riazi, 1997). Regarding the

[6] The stratum statistic (H_p) is derived from the person separation estimate, G_p, and represents the number of strata with a minimum separation of three standard deviation units; specifically, the formula is $H_p = (4G_p + 1)/3$. The basis for using that criterion is that a significance level of $p = .05$ corresponds to a separation difference of 2.79; the value of three is simply the next largest integer.

Table 1
Experience, competence, and anxiety associated with different tasks.

Task	Experience	Competence	Anxiety
Memo	2.81 (1.05)	3.06 (0.74)	2.13 (1.06)
E-mail to teacher	2.08 (0.92)	2.69 (0.79)	3.59 (1.05)
Internet chat	2.14 (1.15)	2.77 (0.95)	2.78 (1.10)
Self-intro	2.99 (0.73)	3.26 (0.73)	2.70 (0.98)
5-Page report	2.39 (1.14)	2.50 (0.86)	3.83 (1.04)
Speech	1.82 (0.88)	2.13 (0.89)	4.17 (0.98)

Note. These results were from a 5-point Likert scale with a mean of 3. Standard deviations are indicated parenthetically.

coding scheme, both males and females strongly preferred direct feedback and coded feedback to indirect or minimalistic feedback (Table 2). The only two significant differences occurred vis-à-vis 'grade only' and 'confirmation only': Although males and females objected to receiving such limited feedback, females objected more strongly.

However, actions taken in response to feedback were minimal, with students generally reading feedback (3.03) and making notes on their work (3.18), yet less frequently revising (2.51) or asking questions about the feedback (2.67). Of the student responses regarding actions taken with feedback, in excess of 90% echoed the sentiments of one student: "I read [the teacher's] comments, looked up meanings I didn't know, and made some notes on my papers, but I never asked questions."

Affective reactions to feedback were positive, with students indicating that they were happy to have received feedback (3.73) and that the feedback increased their motivation (4.12), which in turn encouraged them to exert further effort (3.81); as one student commented, "I am happy when I receive feedback, and it makes me want to work harder." Feedback increased their self-confidence to a smaller extent (3.31), which is not a surprising result. At the same time, they indicated few negative reactions, feeling neither depressed (2.23) nor shocked (2.40) about feedback; moreover, receiving feedback had little effect on whether they liked their teacher (1.53).

Regarding the mode of feedback (Table 3), students also indicated a statistically significant preference for red marking (3.26) to marking with other colors (3.09), which mirrors the preference of males; females, however, had essentially the same preference for red and blue. Students strongly preferred handwritten feedback to computer-generated feedback (3.58–2.49): as noted by one enthusiastic student, "I really like when my teacher writes comments by hand since the computer is not so friendly!" Males, although not supportive of e-feedback (2.60), were more tolerant of it than were females (2.36). Feedback written on the manuscript (3.89) was preferred to feedback provided on a separate paper (2.16). For the respective genders, nearly all differences within sets were significant: Both males and females strongly preferred handwritten feedback to e-feedback and feedback directly on manuscripts to that provided on a separate paper. For ink color, however, males preferred red ink (3.27) to blue ink (2.96), whereas females indicated no preference for one color over the other (3.24 vs. 3.23). All comparisons noted here were statistically significant at the $p < .05$ level after an FDR correction for the set of items.

While no differences were found based on proficiency level or anxiety level, several minor differences were observed based on gender. Females preferred detailed, direct feedback more than males did, while males indicated somewhat higher anxiety concerning feedback. Males were also slightly more receptive to very minimal feedback (only a confirmation mark or only a grade) than were females. Males preferred red feedback, whereas females did not exhibit a preference for one color over another. Females also more strongly disliked computer-generated feedback and feedback provided on a separate paper. In terms of proficiency and anxiety, however, males and females indicated no statistically significant differences.

Table 2
Student feedback preference by gender.

Task	Gender		Effect size[a]
	Female	Male	
Focus of feedback			
Mechanical	4.40 (0.57)	4.26 (0.74)	0.21
Content	4.12 (0.74)	3.96 (0.95)	0.18
Type of feedback			
Explicit (direct)	4.27 (0.77)	4.16 (0.85)	0.13
Coded	4.23 (0.74)	4.11 (0.88)	0.15
Location	2.69 (0.94)	2.92 (1.01)	0.23
Content only	2.39 (1.01)	2.44 (0.98)	0.05
Grade only	1.77 (0.83)**	2.07 (0.99)**	0.33
Confirmation only	1.48 (0.65)**	1.72 (0.86)**	0.32
No feedback	1.19 (0.49)	1.34 (0.73)	0.26

Note. A double asterisk indicates that the difference between the respective means was statistically significant at $p < .01$. Standard deviations are shown in parentheses.

[a] The effect size is Cohen's *d*.

Table 3
Student preferences for mode of feedback.

Task	Gender		Total	Effect size[a]
	Female	Male		
Red pen	3.24 (0.86)	3.27 (0.97)	3.26 (0.92)	0.03
Blue pen	3.23 (0.84)*	2.96 (0.88)*	3.09 (0.88)	0.31
Handwritten	3.66 (0.85)	3.51 (0.86)	3.58 (0.86)	0.18
PC-generated	2.36 (0.79)*	2.60 (0.82)*	2.49 (0.82)	0.31
On manuscript	3.97 (0.69)	3.83 (0.71)	3.89 (0.70)	0.19
On separate paper	2.02 (0.73)*	2.26 (0.86)*	2.16 (0.83)	0.30

Note. A single asterisk indicates that the difference in the respective means was statistically significant at $p < .05$. Standard deviations are shown in parentheses.
[a] The effect size is Cohen's *d*.

Higher levels of proficiency corresponded with lower levels of anxiety across the spectrum of activities queried, which was as expected. Perusal of results from low-proficiency (LP) and high-proficiency (HP) groups yielded some statistically significant differences. HP students were more likely to question the teacher after receiving feedback than were LP students (2.83 vs. 2.52). HP students were also more positively inclined toward affective reactions, indicating that feedback engendered significantly higher levels of happiness and self-confidence. In other areas, however, HP and LP students yielded no statistically significant differences.

Low-anxiety (LA) and high-anxiety (HA) students exhibited more differences. The LA students were more likely to take action when provided with feedback (3.66 vs. 3.13), whereas the HA students were more likely to react negatively in such ways as feeling depressed or shocked while being less happy about having received feedback. Note, however, that those differences were not large, being significant at the FDR-corrected $p < .05$ level: the HA students endorsed feeling happy about feedback (3.62), which was lower than the 3.83 indicated by low-anxiety students. For the mode of feedback (e.g., red ink vs. blue ink and handwritten vs. e-feedback), the means for HA and LA students showed no statistically significant differences.

5. Discussion

The results of the analyses indicated that although students had some experience with the tasks queried, they perceived a relative dearth of competence in those tasks. For the longer and more complex tasks such as writing a 5-page report, the respondents indicated lower levels of proficiency and experience that corresponded with higher levels of anxiety. The modest proficiency and experience levels highlight one shortcoming of secondary education in Japan, in which participants seldom (if ever) write lengthy compositions. This also might reflect a somewhat limited level of competence with word-processing software, a not uncommon reality with Japanese university students (Lockley, 2011; Murray & Blyth, 2011). In the courses taught, considerable time was necessary to enhance the participants' basic ability to manipulate Microsoft Word, and few students were proficient typists.

Second, participants indicated a strong preference for direct, detailed feedback, a finding which corresponds with earlier findings that students prefer to have all surface-level errors corrected to the largest extent possible (e.g., Radecki & Swales, 1988). Furthermore, Timson et al. (1999) found that 1228 EFL learners at nine Japanese universities strongly endorsed the notion that error correction is "necessary and desirable" for increasing L2 proficiency. Indeed, such results contravene that veteran teacher's anti-feedback sentiment and likely reflect the propensity in primary and secondary education in Japan to provide detailed feedback to students. Participants were, however, considerably less enthusiastic about indirect feedback, a finding which suggests its use might be counterproductive in spite of the benefits of negotiating meaning found in previous research (Ellis, Tanaka, & Yamazaki, 1994; Gass & Mackey, 2006; Long, 1996). Moreover, the time invested in addressing indirect feedback could be perceived as an extravagant use of a limited commodity (recall that the participants in this study received a scant 75 minutes of writing instruction per week).

Responses regarding affective reactions to feedback indicated that many students exhibited a well-grounded, mature manner when dealing with feedback. The level of maturity of students has been implicated in success in learning outcomes (Young, 2000), and the even-keeled responses here indicate that corrective feedback is neither problematic for these particular students nor a source of potential concern for the instructor. The color of the feedback was viewed somewhat ambiguously, with only a slight preference for red over blue among males. The primary concern was apparently an eminently practical one far removed from any worries about affect: as noted by one student, "Any color for feedback is OK as long as it's easy to see." Concerns expressed regarding the color of corrective feedback (e.g., Elliot et al., 2007) thus appear to warrant less attention in this context than in such venues as North America.

Regarding the mode of feedback, students indicated a strong preference for handwritten feedback over e-feedback, which is likely a relic of the handwritten correction provided throughout primary and secondary education in Japan. Furthermore, feedback written directly on their manuscript was much preferred to feedback provided on a separate paper.

Actions taken with feedback were informative yet somewhat disappointing. The participants did relatively little with feedback beyond sometimes reading it and making notes on their corrected papers; revisions and asking questions about the

feedback occurred less frequently. Even though one group was required to revise and the other was not, the two groups were statistically indistinguishable regarding actions taken with feedback. This was a somewhat surprising finding because it suggests that there is an underlying level – indeed, a somewhat moderate level – at which students exerted themselves upon receiving feedback. Thus, although feedback was strongly desired, rather minimal action was taken with it regardless of whether revision was required or not. One possible explanation is that students had limited experience using feedback, which suggests that training and practice using feedback would be prudent (Min, 2005, 2006; Ross, 2006).

6. Conclusion

In this study we strove to investigate several issues concerning feedback on writing in EFL classes at one university in Japan. The students queried were found to have more experience and proficiency in certain tasks (e.g., writing a memo) that corresponded with lower levels of anxiety; the reverse was also true with tasks of lower competence and experience associated with higher levels of anxiety. Students also exhibited strong preferences for direct feedback rather than indirect or no feedback.

An unanticipated finding was that students were generally very positive regarding feedback, exhibiting a very mature, well-grounded acceptance of feedback that suggests that negative affective reactions were not a major source of concern in this context. Moreover, handwritten feedback was strongly preferred to e-feedback, a finding that indicates increased use of e-feedback needs to proceed with due attention paid to inculcating students to its use.

The fact that students took very modest action vis-à-vis feedback and that it appeared to be independent of whether feedback was mandatory or not was somewhat surprising. Perhaps indicative of a general student malaise, it highlights the need to train and require students to put feedback to good use. On this point, that veteran teacher appears to have been correct, but the relatively small degree to which feedback is incorporated by students represents not so much a fait accompli as another necessary and teachable skill in the writing classroom.

Finally, the issue of the color of feedback, which has been found to adversely affect cognitive performance in North American contexts, appeared to be of considerably less concern in these contexts. This may be a cultural relic that reflects the widespread use of red for correction in Japan. Females exhibited nearly equal preference for feedback of different colors, while males preferred red feedback. It remains, however, an area in which little research has been attempted and which might be of importance in L2 learning.

The current study thus yielded several kernels of information. Students generally reacted positively to feedback on their EFL writing, yet they exhibited strong preferences for detailed, handwritten feedback that addressed both content and mechanical errors. The color of feedback appears to be an issue of less concern in this context than in North American contexts. Finally, even when feedback is provided, students seem to utilize it only to a modest extent. It is our hope that these findings will inform the work of writing teachers in similar contexts as well as contribute to future research.

Among the possible avenues for future research are closer looks at the students involved, for example through the use of detailed student profiles that include such elements as learning style(s). While the use of self-assessment is common and deemed reliable in some quarters (e.g., Ross, 2006), replicating the current study with more “hard” data in addition to self-assessed proficiency would allay concerns about reliability. Moreover, research on the effect of training students to utilize feedback would further illuminate a lightly-researched area. As noted, research on color exists in such fields as marketing and ergonomics, but there is a dearth of such research in SLA. Given the reality that SLA occurs in myriad contexts and that colors may have different meanings in different cultures (Aslam, 2006), replicating the current study in other contexts would be wise. Finally, replication of this study in the future as e-feedback makes more inroads into evaluation and assessment would also be an appropriate endeavor.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.system.2013.12.023.

References

Aslam, M. M. (2006). Are you selling the right colour? A cross-cultural review of colour as a marketing cue. *Journal of Marketing Communications, 12*(1), 15–30. http://dx.doi.org/10.1080/13527260500247827.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological), 57*(1), 289–300. Retrieved from http://www.jstor.org/stable/2346101.

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing, 17*, 102–118. http://dx.doi.org/10.1016/j.jslw.2007.11.004.

Borg, S. (2003). Teacher cognition in grammar teaching: a literature review. *Language Awareness, 12*, 96–108. http://dx.doi.org/10.1080/09658410308667069.

Bower, J., & Kawaguchi, S. (2011). Negotiation of meaning and corrective feedback in Japanese/English eTandem. *Language Learning & Technology, 15*(1), 41–71. Retrieved 17.07.12, from http://llt.msu.edu/issues/february2011/bowerkawaguchi.pdf.

Brodie, L., & Loch, B. (2009). Annotations with a tablet PC or typed feedback: does it make a difference?. In *Proceedings of the 20th Australasian Association for Engineering Education conference* (pp. 756–770) Adelaide: University of Adelaide.

Burgess, J., & Etherington, S. (2002). Focus on grammatical form: explicit or implicit? *System, 30*(4), 433–458. http://dx.doi.org/10.1016/S0346-251X(02)00048-9.

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in accuracy and fluency of L2 student writing. *Journal of Second Language Writing, 12*, 267–296. http://dx.doi.org/10.1016/S1060-3743(03)00038-9.

Conrad, S., & Goldstein, L. (1999). ESL student revision after teacher written comments: texts, contexts and individuals. *Journal of Second Language Learning, 8*, 147–180. http://dx.doi.org/10.1016/S1060-3743(99)80126-X.
Cotterall, S. (1995). Readiness for autonomy: investigating learner beliefs. *System, 23*(2), 195–205. http://dx.doi.org/10.1016/S0346-251X(95)00008-8.
Denton, P., Madden, J., Roberts, M., & Rowe, P. (2008). Students' response to traditional and computer-assisted formative feedback: a comparative case study. *British Journal of Educational Technology, 39*(3), 486–500. http://dx.doi.org/10.1111/j.1467-8535.2007.00745.x.
Dheram, P. K. (1995). Feedback as a two-bullock cart: a case study of teaching writing. *ELT Journal, 49*(2), 160–168. http://dx.doi.org/10.1093/elt/49.2.160.
Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, New Jersey, USA: Erlbaum.
Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 197–262). Cambridge, UK: Cambridge University Press.
Duppenthaler, P. M. (2002). The effect of three types of written feedback on student motivation. *JALT Journal, 24*(2), 130–151. Retrieved 10.10.10, from http://jalt-publications.org/files/pdf-article/jj-24.2-art2.pdf.
Elliot, A. J., Maier, M. A., Moller, A., Friedman, R., & Meinhardt, J. (2007). Color and psychological functioning: the effect of red on performance in performance attainment. *Journal of Experimental Psychology: General., 136*, 154–168. http://dx.doi.org/10.1037/0096-3445.136.1.154.
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*(3), 353–371. http://dx.doi.org/10.1016/j.system.2008.02.001.
Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning, 44*(3), 449–491. http://dx.doi.org/10.1111/j.1467-1770.1994.tb01114.x.
Fathman, A., & Whalley, E. (1990). Teacher response to student writing: focus on form versus content. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 178–190). Cambridge, UK: Cambridge University Press.
Ferris, D. R. (2002). *Treatment of error in second language student writing*. Ann Arbor, Michigan: University of Michigan Press.
Ferris, D. R., & Roberts, B. (2001). Error feedback in L2 writing classes: how explicit does it need to be? *Journal of Second Language Writing, 10*, 161–184. http://dx.doi.org/10.1016/S1060-3743(01)00039-X.
Gass, S. M., & Mackey, A. (2006). Input, interaction and output: an overview. *AILA Review, 19*, 3–17. http://dx.doi.org/10.1075/aila.19.03gas.
Goldstein, L. M. (2004). Questions and answers about teacher written commentary and student revision: teachers and students working together. *Journal of Second Language Writing, 13*(1), 63–80. http://dx.doi.org/10.1016/j.jslw.2004.04.006.
Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: reconsidering the role of assessment feedback in student learning. *Studies in Higher Education, 27*, 53–64. http://dx.doi.org/10.1080/03075070120099368.
Horwitz, E. (1999). Cultural and situational influences on foreign language learners' beliefs about language learning: a review of BALLI studies. *System, 27*(4), 557–576. http://dx.doi.org/10.1016/S0346-251X(99)00050-0.
James, C. (1998). *Errors in language learning and use: Exploring error analysis*. London: Longman.
Jones, R. (1996). Solving some problems of university education: a case study. In R. Debreceny, & A. Ellis (Eds.), *Proceedings of AusWeb96* (pp. 243–252). Gold Coast, Australia: Southern Cross University Press. Retrieved from http://davidtjones.wordpress.com/publications/solving-some-problems-of-university-education-a-case-study/.
Kalaja, P., & Ferriera Barcelos, A. M. (2003). Introduction. In P. Kalaja, & A. M. Ferriera Barcelos (Eds.), *Beliefs about SLA: New research approaches* (pp. 1–4). Dordrecht, The Netherlands: Kluwer.
Kamimura, T. (2006). Effects of peer feedback on EFL student writers at different levels of English proficiency: a Japanese context. *TESL Canada Journal, 23*(2), 12–39. Retrieved from http://www.teslcanadajournal.ca/index.php/tesl/article/viewFile/53/53.
Kellogg, R. T. (1996). A model of working memory on writing. In C. M. Levy, & S. Ransdell (Eds.), *The science of writing* (pp. 57–71). Mahwah, New Jersey, USA: Erlbaum.
Kellogg, R. T. (1999). Components of working memory in text production. In M. Torrance, & G. Jeffery (Eds.), *The cognitive demands of writing* (pp. 57–71). Mahwah, New Jersey, USA: Erlbaum.
Komura, K. (1999). *Student response to error correction in ESL classrooms* (Master's thesis). Sacramento, California, USA: California State University.
Kubota, M. (2001). Error correction strategies used by learners of Japanese when revising a writing task. *System, 29*(4), 467–480. http://dx.doi.org/10.1016/S0346-251X(01)00026-4.
Kwallek, N., & Lewis, C. M. (1990). Effects of environmental colour on males and females: a red or white or green office. *Applied Ergonomics, 21*(4), 275–278. http://dx.doi.org/10.1016/0003-6870(90)90197-6.
Lalande, J. F. (1982). Reducing composition errors: an experiment. *Modern Language Journal, 66*, 140–149. http://dx.doi.org/10.1111/j.1540-4781.1982.tb06973.x.
Lee, I. (2004). Error correction in L2 secondary writing classrooms: the case of Hong Kong. *Journal of Second Language Writing, 13*(4), 285–312. http://dx.doi.org/10.1016/j.jslw.2004.08.001.
Leki, I. (1991). The preferences of ESL students for error correction in college-level writing classes. *Foreign Language Annals, 24*, 203–218. http://dx.doi.org/10.1111/j.1944-9720.1991.tb00464.x.
Leki, I. (2006). 'You cannot ignore': graduate L2 students' experience of and responses to written feedback practices within their disciplines. In K. Hyland, & F. Hyland (Eds.), *Feedback on second language writing: Contexts and issues* (pp. 266–285). New York: Cambridge University Press.
Leki, I., & Carson, J. G. (1994). Students' perceptions of EAP writing instruction and writing needs across the disciplines. *TESOL Quarterly, 28*(1), 81–101. http://dx.doi.org/10.2307/3587199.
Levine, G. S. (2003). Student and instructor beliefs and attitudes about target language use, first language use, and anxiety: report of a questionnaire study. *Modern Language Journal, 87*(3), 343–364. http://dx.doi.org/10.1111/1540-4781.00194.
Liu, J., & Sadler, R. W. (2003). The effect and affect of peer review in electronic vs. traditional modes on L2 writing. *Journal of English for Academic Purposes, 2*, 219–227. http://dx.doi.org/10.1016/S1475-1585(03)00025-0.
Lockley, T. (2011). Japanese students' experiences of ICT and other technology prior to university: a survey. *The JALT CALL Journal, 7*(1), 93–102. Retrieved from http://journal.jaltcall.org/articles/7_1_Lockley.pdf.
Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie, & T. Bhatia (Eds.), *Handbook of research on second language acquisition* (pp. 413–468). New York: Academic Press.
Madden, T. J., Hewett, K., & Roth, M. S. (2000). Managing images in different cultures: a cross-national study of color meanings and preferences. *Journal of International Marketing, 8*(4), 90–107. Retrieved from http://www.jstor.org/stable/25048831.
Maier, M. A., Elliot, A. J., & Lichtenfeld, S. (2008). Mediation of personality and social psychology bulletin. *Personality and Social Psychology Bulletin, 34*, 1530–1540. http://dx.doi.org/10.1177/0146167208323104.
McCutchen, D. (1996). A capacity theory of writing: working memory in composition. *Educational Psychology Review, 8*(3), 299–325. http://dx.doi.org/10.1007/BF01464076.
Mehta, R., & Zhu, R. (2009). Color on cognitive task performances. *Science, 323*, 1226–1229. http://dx.doi.org/10.1126/science.1169144.
Meyerson, P., & Tryon, W. W. (2003). Validating internet research: a test of the psychometric equivalence of internet and in-person samples. *Behavior Research Methods, Instruments, & Computers, 35*(4), 614–620. http://dx.doi.org/10.3758/BF03195541.
Min, H.-T. (2005). Training students to become successful peer reviewers. *System, 33*(2), 293–308. http://dx.doi.org/10.1016/j.system.11.003.
Min, H.-T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing, 15*(2), 118–141. http://dx.doi.org/10.1016/j.jslw.2006.01.003.
Mori, Y. (1999). Epistemological beliefs and language learning beliefs: what do language learners believe about their learning? *Language Learning, 49*, 377–415. http://dx.doi.org/10.1111/0023-8333.00094.
Murray, A., & Blyth, A. (2011). A survey of Japanese university students' computer literacy levels. *The JALT CALL Journal, 7*(3), 307–318. Retrieved from http://journal.jaltcall.org/articles/7_3_Murray.pdf.

Oxford, R. (1994). Where are we regarding language learning motivation? *Modern Language Journal, 78*(4), 512–514. http://dx.doi.org/10.2307/328589.
Peacock, M. (1999). Beliefs about language learning and their relationship to proficiency. *International Journal of Applied Linguistics, 9*(2), 247–263. http://dx.doi.org/10.1111/j.1473-4192.1999.rb00175.x.
Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: the role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research, 63*(2), 167–199. http://dx.doi.org/10.3102/00346543063002167.
Qi, D. S., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing, 10*(4), 277–303. http://dx.doi.org/10.1016/S1060-3743(01)00046-7.
Radecki, P. M., & Swales, J. M. (1988). ESL student reaction to written comments on their written work. *System, 16*(3), 355–365. http://dx.doi.org/10.1016/0346-251X(88)90078-4.
Rasch, G. (1960). *Studies in mathematical psychology: 1. Probabilistic models for some intelligence and attainment tests.* Oxford, UK: Nielsen and Lydiche.
Riazi, A. (1997). Acquiring disciplinary literacy: a socio-cognitive analysis of text production and learning among Iranian graduate students of education. *Journal of Second Language Writing, 6*(2), 105–137. http://dx.doi.org/10.1016/S1060-3743(97)90030-8.
Rinnert, C., & Kobayashi, H. (2001). Differing perceptions of EFL writing among readers in Japan. *The Modern Language Journal, 85*(2), 189–209. http://dx.doi.org/10.1111/0026-7902.00104.
Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly, 20*(1), 83–95. http://dx.doi.org/10.2307/3586390.
Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research & Evaluation, 11*(10). Retrieved from http://pareonline.net/pdf/v11n10.pdf.
Saito, H. (2008). EFL classroom peer assessment: training effects on rating and commenting. *Language Testing, 25*(4), 553–581. http://dx.doi.org/10.1177/0265532208094276.
Saito, H., & Fujita, T. (2009). Peer-assessing peers' contribution to EFL group presentations. *RELC Journal, 40*(2), 149–171. http://dx.doi.org/10.1177/0033688209105868.
Sakui, K., & Gaies, S. J. (1999). Investigating Japanese learners' beliefs about language learning. *System, 27*(4), 473–492. http://dx.doi.org/10.1016/S0346-251X(99)00046-9.
Sawir, E. (2002). Beliefs about language learning: Indonesian learners' perspectives and some implications for classroom practices. *Australian Journal of Education, 46*, 323–337.
Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129–158. http://dx.doi.org/10.1093/applin/11.2.129.
Schulz, R. A. (2001). Cultural differences in student and teacher perceptions concerning the role of grammar instruction and corrective feedback: USA-Colombia. *Modern Language Journal, 85*(2), 244–258. http://dx.doi.org/10.1111/0026-7902.00107.
Silver, N. C., McCulley, W. L., Chambliss, L. N., Charles, C. M., Smith, A. A., Waddell, W. M., et al. (1988). Sex and racial differences in color and number preferences. *Perceptual and Motor Skills, 66*(1), 295–299. http://dx.doi.org/10.2466/pms.1988.66.1.295.
Soldat, A. S., Sinclair, R. C., & Mark, M. M. (1997). Color as an environmental processing cue: external affective cues can directly affect processing strategy without affecting mood. *Social Cognition, 15*(1), 55–71. http://dx.doi.org/10.1521/soco.1997.15.1.55.
Stone, N. J. (2003). Environmental view and color for a simulated telemarketing task. *Environmental Psychology, 23*(1), 63–78. http://dx.doi.org/10.1016/S0272-4944(02)00107-X.
Timson, S., Grow, A., & Matsuoka, M. (1999). Error correction preferences of second language learners: a Japanese perspective. *JACET Bulletin, 30*, 135–149.
Tono, Y., & Kanatani, K. (1995). EFL learners' proficiency and roles of feedback: towards the most appropriate feedback for EFL writing. *Annual Review of English Language Education in Japan, 6*, 1–11.
Vickerman, P. (2009). Student perspectives on formative peer assessment: an attempt to deepen learning? *Assessment and Evaluation in Higher Education, 34*(2), 221–230. http://dx.doi.org/10.1080/02602930801955986.
Ware, P. D., & Warschauer, M. (2006). Electronic feedback and second language writing. Feedback in second language writing: contexts and issues. In K. Hyland, & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 105–122). New York: Cambridge University Press.
Wiegersma, S., & van der Elst, G. (1988). 'Blue phenomenon': spontaneity or preference? *Perceptual and Motor Skills, 66*(2), 308–310. http://dx.doi.org/10.2466/pms.1988.66.1.308.
Yang, N.-D. (1999). The relationship between EFL learners beliefs and learning strategy use. *System, 27*(4), 515–535. http://dx.doi.org/10.1016/S0346-251X(99)00048-2.
Young, P. (2000). 'I might as well give up': self-esteem and mature students' feelings about feedback on assignments. *Journal of Further and Higher Education, 24*(3), 409–418. http://dx.doi.org/10.1080/030987700750022325.
Zhang, S. (1995). Reexamining the affective advantage of peer feedback in the ESL writing class. *Journal of Second Language Writing, 4*(3), 209–222. http://dx.doi.org/10.1016/1060-3743(95)90010-1.