



Castledown

 OPEN ACCESS

Technology in Language Teaching & Learning

ISSN 2652-1687

<https://www.castledown.com/journals/tltl/>

Technology in Language Teaching & Learning, 5(2), 37–58 (2023)
<https://doi.org/10.29140/tltl.v5n2.911>

Online Teacher Written Corrective Feedback Outside of Class Time: (In)effectiveness and Task Engagement



ANDREW J. BARNES 

Waseda University, Japan
a.barnes@fuji.waseda.jp

Abstract

This study examines a common pedagogical approach whereby students view teacher feedback outside of class time. Facilitated through Moodle, it aimed to investigate not only the effect of written corrective feedback (WCF) on learning outcomes, but importantly, how frequently students review the feedback they receive. The study was conducted at a Japanese Women's University in Tokyo (age: $M = 19.6$, $SD = 0.6$). Over a 15-week semester, students were required to write weekly homework assignments. These were then marked by the researcher and returned via Moodle, with students encouraged to read their feedback before attempting the next assignment. In line with existing research in the field, a no-WCF control group was compared against treatment groups that received indirect WCF with meta-linguistic explanations. Further to this, using an online system such as Moodle allowed the provision of linked grammar resource pages through basic HTML code embedded within the WCF itself. At the end of the semester student texts were analysed to measure improvement. When pre-test and post-tests were compared there were no significant differences within groups or between groups on measures of fluency and accuracy. However, what sets this study apart from others that have investigated student engagement, is the use of online activity logs which in combination with survey data allowed the investigation of actual study habits rather than merely those that are self-reported. The results of which give better insight and valid explanations for the outcomes of the study and call into question the efficacy of take-home feedback.

Keywords: written corrective feedback, take-home feedback, task engagement, Moodle, e-feedback

Introduction

Since the mid-1980s, writing teachers have often been trained to follow a process-approach, collaborating with students across multiple drafts emphasizing ideas and structure, and leaving Written Corrective Feedback (WCF) until the end (Ferris, 2002). Though, it is apparent from the literature and the classroom, that WCF is frequently encountered between drafts or when a teacher would return a piece of work after marking and/or grading it. However, whether a student gives much attention to feedback provided by the teacher on a final draft is open to debate, with Krashen writing in 1984, "The

Copyright: © 2023 Andrew J. Barnes. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are in this paper.

research suggests that feedback is useful when it is done during the writing process, i.e. between drafts. It is not useful when done at the end, i.e. comments and corrections on papers read at home and returned to students” (p. 11). Further to this, it could be argued that merely returning a corrected piece of written work without the need for redrafting would only motivate students to spend little, if any, time on reviewing the corrections. As such, some may question why teachers would persist with less-than-optimal feedback strategies.

Historically, WCF has been seen as a presumed necessity within the arsenal of the L2 teacher—even when this went against the best practice methodologies passed down in teacher training, forcing WCF “underground” (Ferris, 2002, p. xii). Additionally, teachers often use the same techniques that they encountered when they were students themselves, as Kennedy (1991) wrote, “the role models that novice teachers observed while they were children continue to hold tremendous sway. Often, despite their intentions to do otherwise, new teachers teach as they were taught” (p. 16). As time passes, procedure becomes unquestioned and falls into the realms of common-sense teaching methodology. Furthermore, investigations into teacher practice often find mismatches between beliefs of what should be done and actual classroom usage (Lee, 2009; Mao & Crosthwaite, 2019). Irrespective of the why, it holds that there are still many teachers, spending countless hours correcting written work, on the assumption that students will carefully review said corrections outside of class-time.

At present, whether WCF should be implemented for L2 learners at all is still under debate. Started by Truscott (1996) and situated around a simple question of whether or not written grammar correction is effective. This study hopes to add to the body of research in response to Truscott’s question, whilst also attempting to answer one that teachers may have about what happens to those corrected assignments once they are taken home. Currently, possible worst-case answers to this second question range from simply ignoring it—if students are not sat down and forced to look at their feedback (Chandler, 2009)—to even cases of teachers seeing their students toss feedback in the waste-paper basket upon exiting the classroom (Guénette, 2007). Even so, many teachers trust students to review what is taken home, though it may be wishful thinking. What happens outside of the classroom is typically out of the range of what teachers are able to check, therefore, by using online Content Management Systems (CMS) such as Moodle, teachers and researchers can potentially track their student’s activities and engagement within their courses; allowing them to provide feedback on course-work and then know if it is viewed (Parkin et al., 2012; Ziyad, 2016).

According to Ellis in 2009, “No study [had] systematically investigated different approaches to revision” such as “students just given back corrected text” with “no revisions required” (p. 99). This research attempts to help fill this gap by investigating student engagement with, and the effectiveness of, WCF which is read outside of class time, when revisions are not required by the teacher.

Literature Review

There has been a growing interest in meta-analyses of the WCF literature to try and determine patterns of successful practice (Biber et al., 2011; Kang & Han, 2015). Three broad areas of research within the literature can be summarized as: *direct versus indirect* and *metalinguistic explanations* (ME); *focused versus unfocused*; and whether or not to provide feedback within a framework of *drafting*.

Direct feedback is defined simply as when “The teacher provides the student with the correct form” (Ellis, 2009, p. 98), while indirect merely highlighting the error location, and ME being typically either grammatical explanations or hints toward a specific issue. Also, whether to include redrafts (or opportunities for revisions) is another consideration for the teacher (see Ellis, 2009). For the purpose of this study *take-home* feedback can be defined as that which is returned to students to be read at home, with no class time given to reviewing or working with the feedback. Typically, this would include when marked papers are physically returned to a student at the end of class, or even digitally (through other means such as a CMS system) outside of class time.

Comparing approaches to written corrective feedback

A number of studies have examined different styles of WCF individually and/or in combination, focusing on the English article system, with varying success. The addition of ME alongside direct WCF (DWCF) has been effective at improving the accuracy of English articles for intermediate-level learners (Sheen, 2007). Further, the use of error codes increased students' ability to self-edit article errors in a subsequent text (Van Beuningen et al., 2012). In the short-term, even just circling errors can be an effective one-time treatment for article errors. However, when longer-term accuracy is considered, such indirect WCF alone appears to be ineffective, whereas ME retains its efficacy (Bitchener & Knoch, 2010). Therefore, when targeting discrete grammar forms, such as articles, ME proves to be quite effective.

ME provided solely on a separate page, however, has been shown to be an ineffective form of focussed feedback (Shintani et al., 2013). Likewise, the inclusion of error codes in-text may be detrimental if many grammar forms are targeted at once (Chandler, 2003). When using such *comprehensive feedback*¹, direct correction or underlining have both been shown to be more effective for reducing errors when writing new pieces (Chandler, 2003). Furthermore, when the inclusion of revisions is considered, comprehensive DWCF has a greater effect on grammatical accuracy than mere writing practice (WP), whilst comprehensive indirect WCF in the form of error codes does not (Van Beuningen et al., 2012). This is again suggestive that when error codes are used in-text, a more focussed form of WCF is advisable. Therefore, alongside how direct the WCF should be, whether to focus on a small number of grammatical issues requires teacher consideration. The next one being whether to include a drafting cycle or not.

Sheen et al. (2009) compared a group receiving DWCF focussing on articles to one focussing on five grammar forms. In immediate post-tests, the article group showed the greatest improvement, followed by a WP group. However, in delayed post-tests there was no longer a significant difference between either form of WCF and WP. Whilst the result was in favour of focusing on a single form as opposed to many at once, this still raised general questions about WCF effectiveness. In a different approach, Shintani et al. (2013) focussed on two grammar forms at once (English articles, and hypotheticals). They examined the effects of DWCF versus ME (on a separate page). In addition, this study investigated the inclusion of Revisions (+R) which meant further to reading their feedback, +R groups were required to make corrections. In the delayed post-test only DWCF+R outperformed a WP group; groups receiving DWCF, ME, or ME+R all had comparable results to WP at the end of the study. However, there was no mention of alternative activities being given to WP whilst the WCF groups worked with their feedback. Similarly, Chandler (2003) Study 1 is often cited as evidence for redrafts, repeated a cycle where a control group received only Indirect WCF (to be reviewed later) whilst an experimental group completed several tasks: receiving Indirect WCF, they made corrections, then further received DWCF, before both groups moved onto the next writing task. Whilst as a research methodology this has proven problematic (see Chandler 2009; Truscott, 2009), anecdotally, it is a popular WCF approach. That being said, even in a controlled classroom setting, it can be difficult, if not impossible, to truly equalize tasks and account for exactly how students interact with feedback. However, a comparative strength of Van Beuningen et al. (2012) is that an extra practice activity was used for the WP group when other groups were revising drafts. Furthermore, their inclusion of a self-correction group made their findings easier to interpret. As a result, making the strongest case, so far, for the effectiveness of redrafting based on WCF, though, with the limitation that it was a one-shot study.

As to providing WCF outside of a process approach to writing, there appears to be an absence in the recent literature. Older studies, such as Kepner (1991) where a treatment group received comprehensive WCF provided on homework journal writing tasks and made no mention of revisions or class time being spent reviewing feedback, are now excluded from the modern cannon due to using a comparison group that received comments on content rather than a control (see Bitchener & Storch, 2016, Chapter 3). Therefore, there is an apparent need to explore the effects of feedback provided outside of a drafting cycle and over several assignments. Regarding the comprehensiveness of the feedback, Lee (2019) argues that classroom practitioners should move to a more focussed approach to the benefit of both teachers (less time consuming, allowing for a greater focus on written content, creating positive

washback) and students (less intimidating, increasing noticing, comprehension of input, and successful hypothesis testing) (see Bitchener & Storch, 2016, for a cognitive model of WCF within SLA). General themes in the literature also suggest teachers provide more direct forms for lower-level learners, but that indirect forms may be more beneficial for long-term development. However, teachers must also take into account feedback preferences of other key stakeholders in their classroom, namely, the students.

Student perspective of WCF

Often as part of WCF research, questionnaires and/or surveys (Chandler, 2003; Ferris, 2002; Ferris & Roberts, 2001) or interviews (Fazio, 2001) have been used to assess different factors such as motivation, language background, or learner preference.

When feedback preference is investigated, typically error correction is expected by students (Chandler, 2003; Ferris & Roberts, 2001). In response to different styles of WCF in a drafting process, Chandler (2003) found that although participants reported direct corrections were easier to carry over, they indicated a preference for underlining with Error Codes as being more helpful. Interestingly, Chandler noted that this “was not an accurate judgment since the data show that both Correction and Underlining were significantly more effective at reducing error in subsequent writing than Underlining with [Error Codes] was” (p. 287, capitalisation in original). The pre-test questionnaire used by Ferris and Roberts (2001) similarly found that participants preferred underlining with codes.

Focusing more on engagement with WCF, Fazio’s (2001) study involving both French majority-language and minority-language students investigated the effects of grammar correction, comments on content, and a combination of the two. From participant interviews, it was found that of the 17 minority-language students only a single participant who received error correction, and a further four participants from the combination group, reported that they always attended to the feedback; with the “general lack of attentiveness to corrective feedback [being] shared by their francophone counterparts” (p. 245). It is such findings which may explain why no significant difference was found in measures of accuracy within and between the groups over the course of that study.

One final consideration relates to the actual administration and delivery of feedback that students may expect. An ever-growing number of learners come to classrooms equipped with laptop computers, and other mobile devices, and with them arises a need for training in their educational usage (Stockwell, 2019). Furthermore, to meet changing demands, many teachers are now asked to provide course content, resources, and feedback through the CMS systems that their educational setting has invested in. It is for these reasons that WCF practitioners may wish to move away from the hand-written styles investigated in the above studies.

e-Feedback

Electronic feedback, or e-feedback, looks at replacing the traditional pen-and-paper interactions between teacher and student with more digital forms, such as e-mail (Seiffedin & El-Sakka, 2017), or Microsoft Word (Ene & Upton, 2014; Milton, 2006). Studies into which have shown positive results reflective of their pen-and-paper counterparts, as well as some similar drawbacks.

Seiffedin and El-Sakka (2017) investigated a process-approach to providing WCF facilitated through email. First, drafts received indirect WCF; then, second drafts received indirect WCF with error codes; finally, third drafts, DWCF. With ten assignments given over 10 weeks, a no-feedback WP group showed a non-significant improvement in accuracy between pre-test and post-test, whereas the experimental group significantly improved. However, it is unclear whether redrafts replaced assignments or were additional to them. Ene and Upton (2014) conducted a study into learner uptake of WCF in EAP writing classes at an American university. Feedback was provided through Microsoft Word’s comment and track changes functions, taking place between four drafts of two major assignments; each assignment from a consecutive semester’s writing course. They reported an astonishingly successful rate of correct grammar uptake of 95% in the first writing class, although this later dropped significantly to 41.7%. Despite this drop, it may be argued that electronic WCF proved to be a viable means through which to provide feedback to learners.

Other researchers have investigated new ways of leveraging technological affordances to implement WCF. Milton (2006), for example, expounded upon the virtues of a web tool called Mark My Words, which allowed teachers to link their WCF to online resources via a Microsoft Word add-on. It was an eloquent solution for pointing students to various resources as part of the WCF process. Milton's ideas can also be expanded upon to include CMS systems such as Moodle, which can act as digital repositories for curated grammar and writing resources.

Feedback engagement and CMS systems

Feedback engagement has become a growing field of interest, with the driving question of WCF's effectiveness being examined through its mitigating factors. Three main research areas can be summarized as "the ways learners engage with written feedback," "factors that influence student engagement with [it]," and since 2021, investigations into the "impact of pedagogical activities" (Mao & Lee, 2023, p. 6). However, for the current study, the fundamental level of task-engagement under investigation is a more narrowed view of behavioural engagement as "time-on-task or participation" (Philp & Duchesne, 2016, p. 55), that is, how frequently students view take-home feedback based on Moodle server logs.

While not specifically looking at feedback, in a prior ELT study using Moodle server logs, Ziyad (2016) measured course engagement via the number of uploads, downloads, comments, and hours online. Students were enrolled in a writing course that used a process approach; drafts being posted to class forums to elicit peer-feedback. Participants were split into six groups, with data collected through focus group interviews alongside Moodle logs. It was found that attitudes toward Moodle were generally positive, but engagement levels varied between the six groups. Two groups "expressed their acceptance of the online platform and were satisfied by its features" yet "failed to translate such an attitude to active engagement" (p. 328).

Outside of the EFL context there have been general education studies looking at engagement with teacher feedback mediated through technology. Based on student interviews, Parkin et al. (2012) reported the following:

It is conventional wisdom that the return of feedback during face-to-face sessions is good practice. The study found, however, that students placed significant value on the flexibility to receive and read their feedback in private surroundings that the online publication of grades and feedback can offer.

(p. 966)

It requires noting that this study used self-selecting sampling where "those who participate are likely to be more engaged with their learning" (p. 964). Furthermore, using interviews to study engagement relies on "the second-hand reporting of behaviour and captures data after a period of time has lapsed" (p. 965). To that end, the provision of feedback through online systems that learners can view in their own time, can allow them to view feedback when they wish alongside also offering a more reliable picture of when and for how long they are likely to spend on it.

Thus, to investigate how frequently students view their feedback requires a research design fitted to the task, being mindful of the educational and classroom context. Any out-of-class activities that would force students to engage with the feedback, such as a drafting cycle, must be stripped away. Though in doing so, this provides an opportunity to examine an under-investigated approach to take-home WCF, whilst reflecting an at times teaching reality for classroom practitioners.

Method

Research questions

The following research questions were posed for the current study.

RQ1. Does "take-home" focused indirect WCF with ME produce a significant improvement in L2 student writing?

- RQ1a. Of the target grammar?
- RQ1b. Of general grammar accuracy?
- RQ1c. Of written fluency?

RQ2. How do students use their feedback?

- RQ2a. Do they go back and look at it again?
- RQ2b. What are student's impressions of online written feedback?

Setting

A total of 61 female college students aged between 18 and 21 ($M = 19.6$, $SD = 0.6$), from three intact classrooms (herein labelled Class I, II, and III), were recruited from a women's university in Tokyo. Class I, II, and III were English medium classes which followed syllabi based around societal issues. All participants were English majors, although there were differences in English ability both within and between classes. Class I was a first-year class of mixed level. Classes II and III were in their second year; II being approximately lower-intermediate, and III considered beginner. All classes focused on oral communication skills, but students were required to complete written assignments for homework, and gave final in-class presentations as part of their grade. All participants' first language was Japanese with an average 11 years of English education experience. Classes II and III had limited experience studying abroad in the USA (approx. 5-6 months) as part of their degree course, which Class I had yet to complete. The researcher took the role of teaching assistant to the class teacher and was delegated the weekly written assignments. Resultingly, all task design decisions needed to balance both the research objectives and the educational targets of the course.

Test materials

Online learning management system. A Moodle system was used to provide students with online course materials, writing assignments and teacher feedback throughout the course.

Writing Assignments. Students completed a weekly writing assignment in Moodle for homework; questions based on course topics. Class I's syllabus including slightly different topics to those of II and III. Two percent of credit was awarded solely based on completion; 150–250 words being the only requirement. This ensured sufficient data for analysis without biasing students toward written accuracy. Assignments for Weeks 1 and 2 assisted with system training. Weeks 3, 6, 9, and 12 acted as pre-test, mid-test, post-test, and delayed post-test, respectively.

Grammar pages. Grammar pages were created in Moodle to provide tracking data when students accessed them through hyperlinks. Grammar explanations were based on those found in *Business Result* textbooks (Grant et al., 2009, p. 115; Hughes & Naunton, 2007, p.121) and the *LearnEnglish* website (British Council, 2018).

Pre-test survey. Questionnaires were distributed during Week 1's class. Alongside gathering background information on participants, questions used a six-point Likert scale and focussed on student expectations and perceptions of English writing tasks and teacher corrective feedback (Lambert, 2015).

Post-test survey. Online questionnaires were distributed to the participants via the Moodle during Week 15's class. Questions were based on examples given in Ferris (2002) and the pre-test questionnaire found in Ferris and Roberts (2001). Questions focussed on three topics: Moodle, Teacher Feedback, and Writing Assignments. Finally, whilst responses were not anonymous, this potentially allowed for direct comparisons with both the pre-test survey and Moodle Server Logs (see 3.8.3).

Subject assignment method

Class I (mixed level) acted as the Control Group ($n = 28$, $M_{age} = 19.0$, $SD = 0.34$). Participants from Classes II (low-intermediate) and III (beginner) were combined and ranked, from highest to lowest, based on their total combined word count for Week 1's and Week 2's assignments; odd number

ranked participants were assigned to Group A ($n = 17$, $M_{age} = 20.0$, $SD = 0.35$); even number ranked participants to Group B ($n = 16$, $M_{age} = 20.0$, $SD = 0.37$).

The target structures

To tailor the choice of target structures to this particular group of students the written assignments produced in Week 1 were further analyzed for common errors. This influenced the selection of target grammar. Two were selected to meet the needs of the 15-week course:

Target Grammar 1. Determiners and the definite article “the”

Target Grammar 2. To-infinitives

These choices were in alignment with EFL educational recommendations (Nation, 2013) and characteristics of Japanese learners (Ellis et al., 2008; Shirahata, 1991). Whereas many one-shot studies solely focused on the definite article’s *anaphoric usage* (Bitchener & Knoch, 2010; Ellis et al., 2008; Sheen, 2007), this study also included *situational* and *cataphoric usage* to better align with course goals (Yoo, 2009). Similarly, considerations of genre influenced the choice of to-infinitives (Biber & Reppen, 2002).

Feedback style

Written feedback was solely provided by the researcher to avoid issues of interrater reliability (Ferris, 2006). Positive feedback was provided in the form of final teacher comments: encouragement and questions about content; advice on improving structure; and providing models when necessary (Ferris, 2006; Truscott, 1996). Teacher comments were semi-generic in nature to reduce workload, whilst modified to be specific and helpful rather than repetitive for the students (Milton, 2006).

Feedback

Grade	2.00 / 2.00
Graded on	Monday, 14 May 2018, 9:47 AM
Graded by	 

Feedback comments

Countable and uncountable

I think that the Japanese education system for teaching English must change. I have **three reason** for this, firstly we don't get **enough times** to talk, secondly not **all teacher** can speak English well, thirdly there are too **many lesson** on grammar. In my experience, in high school English class we spent **a lot of times** on listening, reading and writing. This means that when we try to speak English we are not confident.

My teachers were excellent at English grammar, but they didn't have much experience outside of Japan, so **a lot of my teacher** did not have good pronunciation. So that means it is difficult to learn good pronunciation.

Though learning grammar is important. When I went to study abroad in Boston **many classmate** had bad grammar, but they were not afraid to speak up. I had the opposite problem. I think confidence is important when learning.

So, I think students in Japan should practice speaking more in class, should have **more conversation times** and our teachers should have experience of living or working abroad a little.

Well Done 

This is a well-structured piece of writing, with a clear introduction, body and conclusion. But maybe you could give some more examples of what you would want to be taught in class. Do you think it is possible to study grammar and speaking at the same time?

Overall, good job.

Please review the below grammar

[Countable and Uncountable Nouns](#)

Figure 1. Dummy example of student WCF used for Moodle system training (Week 2)

Table 1. *Hover Over metalinguistic tags by treatment week*

Treatment	Target Grammar 1 (articles)	Target Grammar 2 (to-verb)
Week 1	<i>Determiners and the definite article “the”</i>	<i>To-infinitives (1)</i>
Week 2	<i>The definite article “the” (1)</i>	<i>To-infinitives (2)</i>
Week 3	<i>The definite article “the” (2)</i>	<i>To-infinitives (3)</i>

Focused WCF was indirect in style; errors were marked-up in red and underlined, embedded with a hyperlink, along with a meta-linguistic tag as a hover effect (Figure 1).

Furthermore, teacher comments ended with a blue hyperlink to that week’s target grammar page. Upon clicking any of the errors marked up in red, or the final link in blue, a grammar page would be opened. These contained English explanations and examples of the target grammar.

Design and Administration

Based on the above choices, the study was quasi-experimental in nature. Following two initial weeks of system training, the WCF treatment schedule (see Appendix A) followed a matrix design for Groups A and B from Assignments 2 through 7. From Assignment 8 onward, WCF stopped. Importantly, both treatments groups and the control group received general feedback in the form of teacher comments for Assignments 1 through 13.

WCF was provided in a three-week sequence from Week 4: Group A focussed on to-infinitives, while Group B focussed on articles. From Week 7, the focus of correction for both groups switched, lasting for a further three weeks. Each week, a new metalinguistic tag (Table 1) and its associated grammar page was introduced, gradually adding complexity.

A research intervention occurred after the mid-test, due to a large drop in students viewing feedback and grammar pages despite verbal reminders: a hyperlink was included on all assignment question pages, from Week 7 onward, instructing students to view their teacher feedback (Figure 2). At the end of Week 6’s class, students were also informed that they would start receiving WCF on a different grammar point for 3 more weeks.

Week 7 Assignment

If you could have any type of artificial intelligent tool, what would you like to have? Describe what it would do and why you would like it.

- Please review your teacher feedback for [Week 5 Assignment](#) before answering the above question.
- Please time yourself, and include how many minutes you took at the end of your writing, e.g. (22 minutes)

Figure 2. *Assignment question with hyperlink added (intervention)*

During the final class (Week 15), students completed the post-test survey, after which, the researcher thanked the students for their participation and dealt with any ad-hoc issues or questions.

Statistical analysis

Selection of Measures. Wolfe-Quintero et al. (1998) reviewed SLA research measures for written fluency, accuracy, and complexity. Table 2 below summarises their findings for best measures for fluency and accuracy. For this study, measures based on t-units were chosen in alignment with research in the field (see Seiffidin & El-Sakka, 2017). Total word count was divided by total number of t-units (W/T) to get a measure of fluency in line with the above table. Furthermore, total number of error free t-units divided by total number of t-units (EFT/T) provided the measure of general accuracy. Due to the low word count of student assignments, the measures CorrIndefArt/T and CorrDefArt/

Table 2. Summary of best developmental measures

Best measures of development so far	
Fluency:	W/T, W/C, W/EFT
Best measures of 'error' so far	
Accuracy:	EFT/T, E/T
Best potential detailed measures	
Accuracy:	CorrIndefArt/T, CorrDefArt/T
Source: Wolfe-Quintero et al., 1998, pp. 119-125	

T were combined to form a new measure for this study: total number of correct articles divided by total T-units, recorded as CorrArt/T. This followed recommendations from Wolfe-Quintero et al. (1998) that ratios should be used when analysing written assignments outside of timed conditions. Finally, similar to that used for articles, a new measure was decided upon for to-infinitive analysis: total number of correct to-infinitives was divided by total number of t-units to give a detailed measure, represented by CorrTo/T.

Student Usage Logs. Moodle server logs were exported into Microsoft Excel. The logs then separated out by individual student and colour coded for (i) teacher actions: providing feedback and grades (red); and (ii) student actions: completing or updating assignments (purple), viewing the screen containing instructor feedback (green), and grammar page visits (yellow). Thus, allowing the researcher to manually check and record subject actions. However, although Moodle included a time stamp for each action, there was no way for the researcher to determine how long a subject truly spent on any one page. Thus, any second access of an assignment feedback page (green) within the same time stamp were discounted under the assumption there were likely page refreshes. However, the number of grammar page views (yellow) could be more easily discerned — each one signifying a click on a hyperlinked correction. Thus, the Moodle server logs could be reduced to binary values reflecting student Behavioural task-engagement.

Interrater reliability. A second teacher was given a 10% sample of the student texts to calculate an interrater reliability score using Pearson's product-moment correlation coefficient (r). For the total number of t-units, the correlation was $r = .984$; for accurate t-units, it was $r = .766$, for incorrect t-units, it was $r = .761$, for article accuracy, it was $r = .986$, and for to-infinitive accuracy it was $r = .868$.

Results

Pre-test survey

The pre-test survey was completed by Group A ($n = 16$), Group B ($n = 17$), and Control ($n = 27$) (see Appendix B for detailed results). First, examining the teacher's role, there was a general agreement that teachers should rewrite (96.67%) and explain (100.00%) grammar errors. Further to this, looking at structural errors, reveals another general expectation that errors should be rewritten (96.61%) and explained (94.92%). One-way ANOVA ($F(2, 57) = 2.631, p = .081$) revealed a significant difference at $p < .1$ on whether teachers should rewrite grammar errors. Though all three groups were in general agreement that teachers should, Group B displayed stronger agreement than Group A and the control group (see Appendix B, Table B2).

The next section of the survey looked at the role of the student. Of the participants, 93.2% agreed with the statement that they always review teacher feedback, with 35.6% stating they strongly agree. Further to this 93.2% agreed with the statement that they always review corrections provided by their teacher, with 33.90% displaying strong agreement. 84.8% of participants agreed with the statement that they review grammar topics they are weak at. Finally, 83.1% agreed that they review structure topics they are weak at. These responses displayed general acknowledgement of the student role when teacher feedback is received.

Comparability of groups before treatment

For comparability of results, only students that had completed 12 assignments were included: Group A ($n = 9$), Group B ($n = 10$), and the control group ($n = 12$). This halved the potential dataset from 61 participants to 31. One-way ANOVA was then conducted on the student baseline data (pre-test, Week 3) to discern if there were any significant differences between the three groups in fluency, general accuracy, article accuracy, and to-infinitive accuracy:

- Fluency (WC/T): $F(2,28) = .285, p = .754$. The result was *not* significant at $p < .05$.
- General accuracy (EFT/T): $F(2,28) = .168, p = .846$. The result was *not* significant at $p < .05$.
- Article accuracy (CorrArt/T): $F(2,28) = .508, p = .607$. The result was *not* significant at $p < .05$.
- To-infinitive accuracy (CorrTo/T): $F(2,28) = 2.39, p = .110$. The result was *not* significant at $p < .05$.

This showed that there were no initial differences between all three at the point of the pre-test.

Descriptive statistics across time

One-way ANOVA for repeated measures were conducted across the four sample points (pre-, mid-, post-, and delayed post-test). Paired-sample t-tests were then conducted as necessary using pre-test, and delayed post-test results. Results were found to be *not* significant at $p < .05$ for fluency, general accuracy, article accuracy and to-infinitive accuracy for Groups A and B, and the control group. Though non-significant, the control group showed increases across all four measures at the delayed post-test, whereas both treatments groups showed decreases. This provided further support that the administration of WCF in this study, was ineffective across time (see Tables 3 through 6 below for the means and standard deviations for each measure). Also included are the effect sizes across time, each subsequent test compared against the respective pre-test result—a correction factor was included in the calculation of *Cohen's d* to account for the small sample sizes.

Comparability of groups in the delayed post-test

One-way ANOVA was conducted on the 31 remaining students' delayed post-test data to discern if there were any significant long-term difference between the three groups in:

- Fluency (WC/T): $F(2,28) = .037, p = .963$. The result was *not* significant at $p < .05$.
- General accuracy (EFT/T): $F(2,28) = .622, p = .544$. The result was *not* significant at $p < .05$.
- Article accuracy (CorrArt/T): $F(2,28) = 2.16, p = .135$. The result was *not* significant at $p < .05$.
- To-infinitive accuracy (CorrTo/T): $F(2,28) = 1.33, p = .282$. The result was *not* significant at $p < .05$.

This shows no differences between all three groups at the point of delayed post-test.

Post-test survey

In contrast to dwindling task completion rates, the majority of students completed the post-test survey: Group A ($n = 16$), Group B ($n = 13$), and the control group ($n = 26$). Students were asked about their thoughts on a variety of topics under three distinct headings: Moodle, teacher feedback, and writing assignments (see Appendix C for detailed results).

Focusing on the Moodle's usability, all but two students reported that they understood how to view the teacher feedback, both being in the control group. Further to this, one student, who was in Group B, reported that they had problems viewing the feedback on their system. In conversation, the student explained that it was only until the late stages of the experiment that she discovered the feedback area, which indicates a possible issue with student training rather than an issue with the Moodle software. This was also reflected in two students not understanding how to view the grammar pages.

The second section covered the frequency that students viewed and interacted with their feedback. In

Groups	<i>n</i>	Pre-test			Mid-test			Post-test			Delayed Post-test		
		<i>M</i>	<i>SD</i>	<i>Cohen's d</i>	<i>M</i>	<i>SD</i>	<i>Cohen's d</i>	<i>M</i>	<i>SD</i>	<i>Cohen's d</i>	<i>M</i>	<i>SD</i>	<i>Cohen's d</i>
Group A	9	12.089	1.62		10.076	1.117	-1.134	11.147	1.188	-0.520	11.768	2.333	-0.125
Group B	10	12.57	1.98		10.035	1.405	-1.193	10.828	1.437	-0.814	11.981	2.059	-0.236
Control	12	11.818	2.968		11.254	2.057	-0.186	11.393	2.294	-0.135	12.042	2.549	0.068

Groups	<i>n</i>	Pre-test		Mid-test			Post-test			Delayed Post-test		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>Cohen's d</i>	<i>M</i>	<i>SD</i>	<i>Cohen's d</i>	<i>M</i>	<i>SD</i>	<i>Cohen's d</i>
Group A	9	0.416	0.153	0.344	0.124	-0.405	0.388	0.141	-0.149	0.323	0.206	-0.402
Group B	10	0.453	0.164	0.377	0.184	-0.352	0.372	0.197	-0.361	0.366	0.178	-0.411
Control	12	0.415	0.183	0.425	0.169	0.048	0.443	0.128	0.149	0.42	0.211	0.021

Groups	n	Pre-test		Mid-test			Post-test			Delayed Post-test		
		M	SD	M	SD	Cohen's d	M	SD	Cohen's d	M	SD	Cohen's d
Group A	9	0.53	0.373	0.335	0.169	-0.528	0.322	0.19	-0.551	0.312	0.263	-0.530
Group B	10	0.583	0.15	0.358	0.127	-1.308	0.376	0.209	-0.919	0.405	0.234	-0.732
Control	12	0.447	0.37	0.624	0.382	0.397	0.407	0.228	-0.110	0.513	0.171	0.193

Groups	n	Pre-test		Mid-test			Post-test			Delayed Post-test		
		M	SD	M	SD	Cohen's d	M	SD	Cohen's d	M	SD	Cohen's d
Group A	9	0.438	0.223	0.168	0.066	-1.287	0.242	0.116	-0.864	0.288	0.196	-0.560
Group B	10	0.446	0.186	0.141	0.098	-1.657	0.27	0.107	-0.937	0.322	0.131	-0.623
Control	12	0.265	0.242	0.094	0.068	-0.811	0.25	0.189	-0.058	0.408	0.192	0.552

The final section looked at the subject of instructor feedback and questions related to English gram-

Table 7. *Writing assignments*

	Group A (n = 16)	Group B (n = 13)	Control (n = 26)
18. Did you use a word processor, such as Microsoft Word to help you write these assignments.			
Yes	25%	69%	38%
No	75%	31%	62%
21. Did you find the writing assignments useful?			
Yes	100%	100%	100%
No	0%	0%	0%

mar (though questions addressing the provided WCF excluded the Control). All but one of the students responded positively that both grammar correction and grammar pages on articles and to-infinitives were helpful to them. It should be noted that a single student from Group B answered *no* to all questions related to feedback being helpful; understandably, this was the same student that reported having had issues accessing the feedback.

Providing further insight as to how students may have interacted with the WCF, a fairly large percentage across all three groups declared *[English grammar] is not a serious problem for me*. This might be the reason that such participants from Group A (38%) and Group B (31%) did not fully engage with the WCF.

A further question looked at student preference of feedback types. What immediately stood out is that the large majority of students in both Group A (81%) and Group B (77%) chose *Circle my errors and tell me what type of errors they are (verb tense, word choice, etc.)* which is in-line with the style of WCF they received in this experiment. This is interesting when compared to students in the control group, who gave a wide variety of responses. This may imply that students are likely to be influenced by the methods that the teacher chooses to use.

The final set of questions looked at how the students approached the written assignments in general (Table 7). Differing numbers of students across groups had not typed directly into Moodle, instead having used a word processor. As an addendum, these students explained what supplemental tools they may have used. It was found that all but one of these students had used Microsoft Word – that student's response being simply "computer"; in addition to Word another student had used Grammarly (a popular online grammar checking website), and finally one student also mentioned the usage of their memo-pad on their smart phone. Obviously, such admissions prove problematic when assessing the true abilities of students based solely on their texts.

Finally, it is a positive outcome that all students in this study reported that the written assignments had been useful to them. Though some could argue this is merely evidence of acquiescence bias.

Student usage logs

Moodle server log analysis focussed on the 19 subjects, receiving WCF, that had completed 12 assignments: Group A (n = 9); Group B (n = 10).

Teacher feedback. Upon server log analysis of the remaining 19 subjects, it was discovered that only two subjects from Group A and four subjects from Group B followed the expected procedure—checking their feedback (with or without accessing grammar page views) for each of the consecutive six treatment weeks. At the other end of the scale, it was found that two subjects from Group A and three subjects from Group B checked their feedback (with or without accessing grammar pages) once per three treatment phases (twice in total). This accounted for 11 of 19 subjects. The remaining eight were sporadic in usage, including instances of checking the feedback and grammar pages outside of the treatment phase.

Grammar pages. Only one of the 19 students, across Groups A and B, didn't make use of the grammar pages at all. Though, it should be noted this subject only received WCF on one article error in

Table 8. Grammar page views by class and group

Students	n	To-infinitives			Articles		
		Page 1	Page 2	Page 3	Page 1	Page 2	Page 3
Total	33	67%	33%	18%	76%	21%	27%
Class II	19	58%	32%	16%	79%	26%	26%
Class III	14	79%	36%	21%	71%	14%	29%
Group A	17	82%	35%	6%	65%	29%	35%
Group B	16	44%	31%	31%	88%	13%	19%

Week 2, a further single to-infinitive error in Week 6, and one more in week 7. Therefore, due to this unusually high level of accuracy with the target grammar, it could be argued that this student simply did not receive WCF. However, the student could have taken advantage of this positive feedback to test their existing hypotheses about L2 grammar.

Taking a broader view, in Table 8, it can be seen what percentage of the combined original 33 subjects viewed each grammar page at least once. Minor differences in these results suggest that the class level had little impact on use of the grammar page overall.

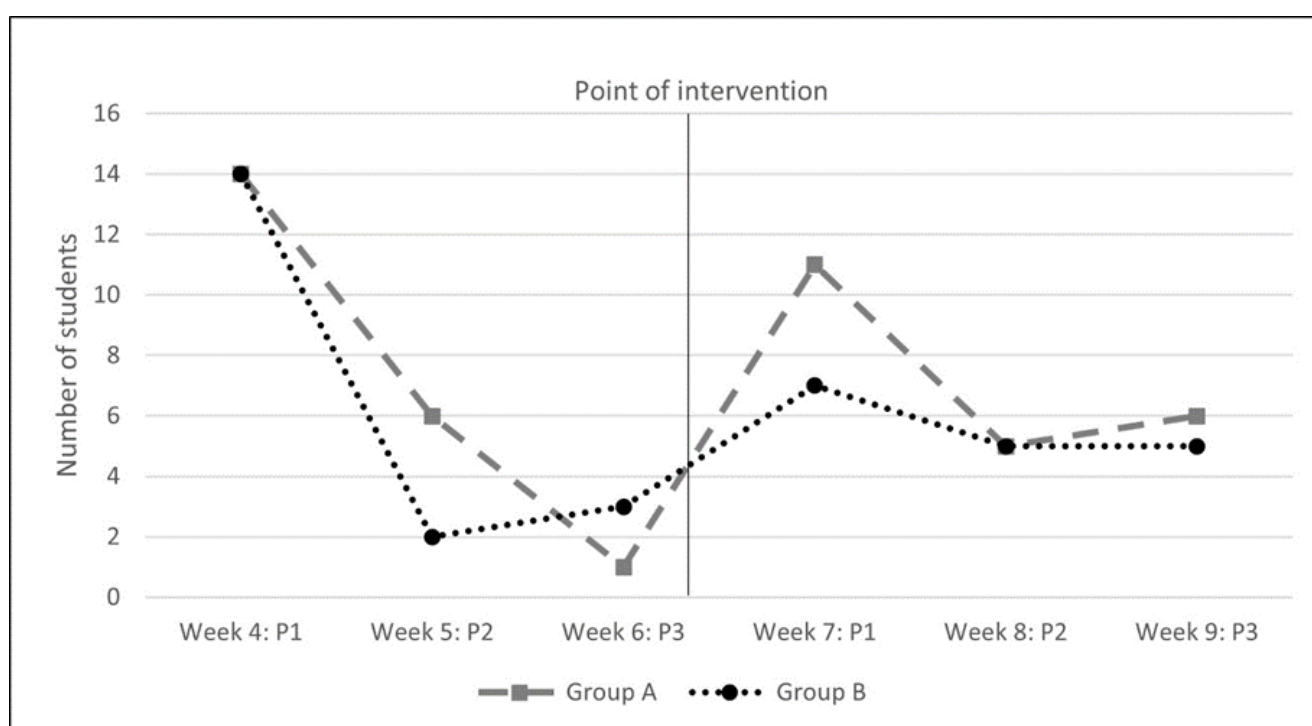
**Figure 3.** Effects of intervention on student numbers viewing grammar pages

Figure 3 displays the potential impact of the researcher intervention, where, within Moodle, students were reminded to check their feedback before starting the next assignment (see Figure 2). The results suggest that along with the change of WCF focus, it had only a small effect on student grammar page interest beyond Page 1.

Finally, of the 19 subjects, 13 viewed one or more feedback pages at a later date (i.e., after the initial

treatment week); of those 13 however, only three viewed the WCF after the six-week treatment phase had passed.

Summary

At first glance, the provision of WCF on articles or to-infinitives had no discernible effect on student progress. However, an examination of the Moodle server logs revealed that, rather than the study being evidence of the ineffectiveness of indirect WCF per se, students were simply not engaging with (or at times even viewing) the feedback. Notably, the lack of task engagement across the classes was at odds with what many reported in their questionnaires. These are important points to be addressed in the next section.

Discussion

Research Question 1: Indirect WCF & accuracy

RQ1 asked whether focused indirect WCF produced a significant improvement in accuracy in the writing of students in an EFL environment. The results were conclusive in so far that the “take-home” WCF did not have a positive effect on the measured outcomes. This ran counter to the significant results for WCF found by other studies that focused on articles (Bitchener & Knoch, 2010; Ellis et al., 2008; Sheen, 2007). However, as with many WCF studies, as Guénette (2007) stated “dissimilarities in methodology make it extremely difficult to compare the results” (p. 42), though this discussion will endeavour to do so.

In terms of article accuracy (RQ1a), and in contrast to the present study, Bitchener and Knoch (2010) found focused ME significantly outperformed WP, displaying significant long-term gains whereas the group receiving only indirect WCF did not. This contrast requires some consideration. Potentially, differences could be due to the range of errors they focused upon and the associated ME being comparatively simple when weighed against the growing complexity of this study’s grammar pages. In addition, the fact that the grammar pages were provided in the target language may have been a further hurdle for the participants in Class III who had lower proficiency. Also, much like the results from Shintani et al. (2013), ME was provided on a separate page which gave a similar end result.

Approaching this issue with data from the Moodle server logs, it was shown that only a total of five students from Groups A and B (15.2%) visited the three different article pages. This means that the majority may have merely perceived their WCF as being indirect in the form of a red underline if they didn’t activate the hyperlink or metalinguistic hover over effect. Alternatively, if they had seen the meta-linguistic tag, they may have assumed that the explanation was the same as a previous week. Unfortunately, the original idea of the hover over was in part so that students could choose to use them or not when it came to the detailed ME pages—it appears many chose the latter. Taking this into account, it could be argued as a failure at the point of detection (Tomlin & Villa, 1994) for why the long-term results were closer to Bitchener and Knoch’s (2010) indirect feedback group and not the ones that received ME.

In terms of general grammar accuracy (RQ1b), the findings appear to be in-line with those of early studies that claimed WCF as grammar correction was ineffective in the long-term (Truscott, 1996) and supports the efficacy of WP (Fazio, 2001; Sheen et al., 2009). In contrast, it was Chandler (2003) Study 1 that demonstrated WCF to be superior to WP in the long-term. Although this may have been due to time-on-task differences between Chandler’s groups, and that class time was given to revisions.

A surface conclusion may be that the current study appears to support research demonstrating the ineffectiveness of WCF. However, through the inclusion of Moodle server logs, this outcome does not appear to be as definitive as first thought. Chandler’s (2009) assertion may therefore be correct that if the feedback is not salient, it in effect is not received (Schmidt, 1995). The lack of engagement with the feedback in this study attests to this. Resultantly, this points to a more specific finding that it is “take-home” WCF that is ineffective, not WCF itself.

Research Question 2: Student use of feedback

RQ2a asked whether students revisit older feedback multiple times. It was found that this question was somewhat idealistic, a more pertinent question may have been whether or not students viewed take-home feedback at all. Although most students reported in the survey that they did indeed check their teacher feedback, an analysis of the Moodle server logs showed this not to be the case. Firstly, it can be seen that once the participants were limited to just those that actually completed all the assignments ($n = 19$)—who it could be presupposed were amongst the most motivated students (Gardner, 2006)—that only a third of the students from Groups A and B combined checked their feedback each week. For WCF this was a total of six students, which it should be noted was four less than those claimed to have done so. Such findings were strikingly similar to those of Fazio's (2001) interviews with French language students in Quebec. However, addressing RQ2a directly, it can be said that a little over half of the remaining 19 participants (ten students) across the treatment groups did indeed go back to view feedback from earlier assignments as the semester progressed. This is a positive result in that students were clearly attaching value to the feedback to warrant rereading it, though of course on the proviso they had read it in the first place. It cannot be ignored though, that the general findings at large were reminiscent of Guénette's (2007) anecdote of seeing their students toss their corrected writings in the wastepaper basket as they left for home.

This in turn starts to address RQ2b, which asked what students' impressions of online written feedback were. In addition to Moodle server logs showing that students did view older feedback for a second or third time, the post-survey revealed that they held a positive view of the style of WCF they received along with the provision of the grammar pages. It could be claimed that results were not due to any perceived lack of efficacy of the WCF from the perspective of the students—participants in Groups A and B almost universally agreed that the provided grammar correction on to-infinitives and articles was helpful to them. In addition, a large majority stated the form of feedback received was also desirable. On the other hand, it could be argued that participants were merely trying to please the researcher (see Weber & Cook, 1972). Upon comparison between the Moodle logs and the post-test survey, evidently, there seems to be a difference between what students say they do and what they actually do.

Student perspective. First it is important to consider student responses to the post-test survey and (in conjunction with the lesson logs) what this might mean for the surface validity of the questions that this study mimicked (Ferris, 2002; Ferris & Roberts, 2001). Students generally reported that the feedback was in a style they were happy with, they had no problems with the system, and universally that both the written assignments and feedback were useful to them. Furthermore, every indicator of external motivating factors which the researcher had control over pointed toward engagement with the course. This raises the question as to why so many students missed assignments and did not view instructor feedback.

The fact that the positive views above were not reflected in the Moodle logs adds weight to arguments that what subjects say to researchers may not always be the truth of the matter. This could possibly be due to certain demand characteristics within the experiment which students became aware of and the resultant roles they took on (Weber & Cook, 1972). Although efforts were made to mask the exact purpose of the study, it cannot be guaranteed that this was entirely successful, with many students being at least able to guess its focus based on the post-test survey. It is conceivable that the participants, therefore, took on the role of model students and merely provided the researcher answers which they predicted were most desirable. The near totality of positive response to the assignments and feedback in the post-test survey, may also suggest that there were cultural aspects at play, resulting in these Japanese learners taking on the above role on mass, which might not have been the case in a different educational setting.

Another interesting outcome was a change in feedback preference between pre- and post-test surveys. In the latter, students showed a preference for the style of WCF they had received. The control groups selected a far wider range of preferences, whereas a majority in each of the two treatments groups indicated that they felt that indirect WCF was best. This seems to echo observations from a

review of the literature by Ferris (2002), who writes, “In general, there appeared to be a good match between what students said they wanted in teacher feedback and the types of feedback they reported receiving from teachers” (p. 31). As a result, this points to the students being affable, and willing to change the expectations they placed on the teacher. Still, whatever the reasons, it cannot be ignored that the reliability of the survey is put in a different light based on the reality of the Moodle logs. The insights given are of interest but must take a back seat to what the students actually did with the feedback and available grammar pages.

Teacher perspective. With students looking at teacher feedback intermittently at best, and as a result not showing the expected signs of improvement after consecutive weeks of WCF, then researchers and educators must ask themselves what the impact is, not only on the student, but also on the teacher. Teachers who have laboriously provided feedback for their students may find that as student motivation declines so might their own. Though not recorded to the detail of Chandler (2003), the researcher found themselves spending roughly ten hours a week on marking and providing the feedback to the three groups, and found themselves questioning the merit of spending hours correcting student work (Guénette, 2007). Furthermore, it was disheartening to learn after the fact that most of that time appeared to be wasted, with only the motivated few reading their feedback regularly.

Limitations of the study

Due to this study being conducted across a 15-week semester with assignments completed outside of class time, what started as a fair sample size was affected by the students’ diligence. The number of participants, falling as students failed to complete required tasks, had a bearing on the validity of the descriptive statistics. However, rather than shifting blame onto the students, this speaks to issues of motivation contingent on both the task design and method of feedback under investigation.

Conclusion

The purpose of the study was to determine the effectiveness of WCF supplied through a CMS to Japanese university students when compared against the attainment levels of a WP group. To look at the results of the study without the insights gained from the Moodle server logs, it would be easy to claim that indirect WCF with ME targeting to-infinitives and articles was ineffective, thus reinforcing the argument put forward by Truscott (1996). However, these logs showed that most students were not actually reading their feedback, let alone looking at the grammar pages. As such, the initially perceived outcomes were not the full story, and the Moodle server logs provided realistic insights into how students used their feedback, as well as confirmation and validation of questionnaire responses. On the whole students reported positive views of the process and yet failed to engage with the feedback. Therefore, a case could be made that if said outcomes are weighed against the countless hours spent by the teacher, providing take-home WCF should be viewed as a methodological failure.

Though this research had several issues regarding the validity of the data, it does show some things that can be taken away on a pedagogical level. In studies with more positive results of other WCF research, class time was provided for the feedback thus ensuring that the students actually looked at teacher feedback. An alternative to this may be to require homework redrafts, which would require students to review any WCF in order to edit texts for resubmission. Even so, the researcher cautions against the continued practice of providing students with feedback to be viewed at home, which students may never even look at. In addition to this, given the time constraints for both teachers and students, the workload associated with WCF may have a direct negative impact on the motivation of not only the learner, but the teacher as well, with both expressing frustration at time being ill-spent. Finally, as a positive point, although this study had a negative outcome in terms of the effectiveness of WCF, it did demonstrate how Moodle can be a useful tool for both teachers and researchers, opening up future avenues to explore the effects of WCF on the student.

Although other studies into online feedback have demonstrated its efficacy, replication studies could be performed within Moodle to dissuade any fears that the technology was the reason for the results of the current study. DWCF or the inclusion of meta-linguistic information displayed directly on the feedback page would potentially resolve any issues students may have had with the feedback not

being salient. Furthermore, such a study could test claims made in this paper about take-home feedback.

Final thoughts

Although this is not a call to do away with WCF in line with Truscott (1996), its implementation in true classroom settings requires careful consideration. If teachers are expected to spend countless hours out of class on this undertaking, the question moves toward how to ensure that the students read the feedback without spending valuable class time on it. This is a question not only for the EFL researcher but has implications for pedagogical practice at large.

In the context of writing education, one answer to this question is the inclusion of redrafts. This would necessitate students looking at the feedback outside of class time to make the necessary revisions. However, an issue raised by modern computing, is the temptation for the less industrious students to lean heavily on the word processor at best, and artificial intelligence systems at worst. Thus, a solution which remains is to bring everything back to the classroom; if WCF is to be provided, students need to be assigned class-time toward reviewing it. Should this be done, rewriting of previous assignments may not be necessary. Without further investigation, the solution is uncertain, but this research suggests that the provision of take-home WCF with the expectation that students will carefully review it by themselves is a practice that must be seriously reconsidered.

References

- Biber, D., Nekrasova, T., & Horn, B. (2011). The effectiveness of feedback for L1-English and L2-writing development: A meta-analysis. *ETS Research Report Series*, 2011(1), i–99. <https://doi.org/10.1002/j.2333-8504.2011.tb02241.x>
- Biber, D., & Reppen, R. (2002). What does frequency have to do with grammar teaching? *Studies in Second Language Acquisition*, 24(2), 199–208. <https://doi.org/10.1017/S0272263102002048>
- Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, 19(4), 207–217. <https://doi.org/10.1016/j.jslw.2010.10.002>
- Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development*. Multilingual Matters.
- British Council. (2018). *English grammar*. <http://learnenglish.britishcouncil.org/en/english-grammar>
- 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(3), 267–296. [https://doi.org/10.1016/S1060-3743\(03\)00038-9](https://doi.org/10.1016/S1060-3743(03)00038-9)
- Chandler, J. (2009). Dialogue: Response to Truscott. *Journal of Second Language Writing*, 18(1), 57–58. <https://doi.org/10.1016/j.jslw.2008.09.002>
- Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, 63(2), 97–107. <https://doi.org/10.1093/elt/ccn023>
- 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. <https://doi.org/10.1016/j.system.2008.02.001>
- Ene, E., & Upton, T. A. (2014). Learner uptake of teacher electronic feedback in ESL composition. *System*, 46, 80–95. <https://doi.org/10.1016/j.system.2014.07.011>
- Fazio, L. L. (2001). The effect of corrections and commentaries on the journal writing accuracy of minority-and majority-language students. *Journal of Second Language Writing*, 10(4), 235–249. [https://doi.org/10.1016/S1060-3743\(01\)00042-X](https://doi.org/10.1016/S1060-3743(01)00042-X)
- Ferris, D. (2002). *Treatment of error in second language student writing*. University of Michigan Press.
- Ferris, D. (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: Contexts and issues* (pp. 81–104). Cambridge University Press.
- Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, 10(3), 161–184. [https://doi.org/10.1016/S1060-3743\(01\)00039-X](https://doi.org/10.1016/S1060-3743(01)00039-X)

- Gardner, R. C. (2006). The socio-educational model of second language acquisition: A research paradigm. *EUROSLA Yearbook*, 6(1), 237–260. <https://doi.org/10.1075/eurosla.6.14gar>
- Grant, G., Hudson, J., & McLarty, R. (2009). *Business result: Pre-intermediate student's book*. Oxford University Press.
- Guénette, D. (2007). Is feedback pedagogically correct?: Research design issues in studies of feedback on writing. *Journal of Second Language Writing*, 16(1), 40–53. <https://doi.org/10.1016/j.jslw.2007.01.001>
- Hughes, J. & Naunton, J. (2007). *Business result: Intermediate student's book*. Oxford University Press.
- Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *The Modern Language Journal*, 99(1), 1–18. <https://doi.org/10.1111/modl.12189>
- Krashen, S. D. (1984). *Writing, research, theory, and applications*. Pergamon Press.
- Kennedy, M. M. (1991). Some surprising findings on how teachers learn to teach. *Educational leadership*, 49(3), 14–17. https://files.ascd.org/staticfiles/ascd/pdf/journals/ed_lead/el_199111_kennedy.pdf
- Kepner, C. G. (1991). An experiment in the relationship of types of written feedback to the development of second-language writing skills. *The Modern Language Journal*, 75(3), 305–313. <https://doi.org/10.2307/328724>
- Lambert, J. (2015, December 10). Written corrective feedback: effective for students, Timesaving for teachers. *Cambridge University Press ELT* [Video]. YouTube. <https://youtu.be/AZz3Td-Hc4k>
- Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *ELT Journal*, 63(1), 13–22. <https://doi.org/10.1093/elt/ccn010>
- Lee, I. (2019). Teacher written corrective feedback: Less is more. *Language Teaching*, 52(4), 524–536. <https://doi.org/10.1017/S0261444819000247>
- Mao, S. S., & Crosthwaite, P. (2019). Investigating written corrective feedback: (Mis)alignment of teachers' beliefs and practice. *Journal of Second Language Writing*, 45, 46–60. <https://doi.org/10.1016/j.jslw.2019.05.004>
- Mao, Z., & Lee, I. (2023). Student engagement with written feedback: Critical issues and way forward. *RELC Journal*, 54(1), 1–9. <https://doi.org/10.1177/00336882221150811>
- Milton, J. (2006). Resource-rich web-based feedback: Helping learners become independent writers. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 123–139). Cambridge University Press.
- Nation, P. (2013). *What should every EFL teacher know?* Compass Publishing.
- Parkin, H. J., Hepplestone, S., Holden, G., Irwin, B., & Thorpe, L. (2012). A role for technology in enhancing students' engagement with feedback. *Assessment & Evaluation in Higher Education*, 37(8), 963–973. <https://doi.org/10.1080/02602938.2011.592934>
- Philp, J., & Duchesne, S. (2016). Exploring engagement in tasks in the language classroom. *Annual Review of Applied Linguistics*, 36, 50–72. <https://doi.org/10.1017/S0267190515000094>
- Seiffedin, A. H., & El-Sakka, S. M. F. (2017). The impact of direct-indirect corrective e-feedback on EFL students' writing accuracy. *Theory and Practice in Language Studies*, 7(3), 166–175. <https://doi.org/10.17507/tpls.0703.02>
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1–63). University of Hawai'i Press.
- Sheen, Y. (2007) The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41(2), 255–283. <https://doi.org/10.1002/j.1545-7249.2007.tb00059.x>
- Sheen, Y., Wright, D., & Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System*, 37(4), 556–569. <https://doi.org/10.1016/j.system.2009.09.002>
- Shintani, N., Ellis, R., & Suzuki, W. (2013). Effects of written feedback and revision on learners' accuracy in using two English grammatical structures. *Language Learning*, 64(1), 103–131. <https://doi.org/10.1111/lang.12029>
- Shirahata, T. (1991). The acquisition of English infinitive and gerund complements by Japanese EFL learners (second language acquisition theory & analysis of scholastic abilities). *ARELE: Annual Review of English Language Education in Japan*, 2, 41–50. https://doi.org/10.20581/arele.2.0_41

- Stockwell, G. (2019). Insights from replication on the factors affecting task engagement in mobile-based learning activities. *Technology in Language Teaching & Learning*, 1(1), 33–51. <https://doi.org/10.29140/tltl.v1n1.152>
- Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16(2), 183–203. <https://doi.org/10.1017/S0272263100012870>
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46(2), 327–369. <https://doi.org/10.1111/j.1467-1770.1996.tb01238.x>
- Truscott, J. (2009). Dialogue: Arguments and appearances: A response to Chandler. *Journal of Second Language Writing*, 18(1), 59–60. <https://doi.org/10.1016/j.jslw.2008.09.001>
- Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language learning*, 62(1), 1–41. <https://doi.org/10.1111/j.1467-9922.2011.00674.x>
- Weber, S. J., & Cook, T. D. (1972). Subject effects in laboratory research: An examination of subject roles, demand characteristics, and valid inference. *Psychological Bulletin*, 77(4), 273–295. <https://doi.org/10.1037/h0032351>
- Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity*. University of Hawai'i Press.
- Yoo, I. W. (2009). The English definite article: What ESL/EFL grammars say and what corpus findings show. *Journal of English for Academic Purposes*, 8(4), 267–278. <https://doi.org/10.1016/j.jeap.2009.07.004>
- Ziyad, H. (2016). Technology-mediated ELT writing: Acceptance and engagement in an online Moodle course. *Contemporary Education Technology*, 7(4), 314–330. <https://dergipark.org.tr/en/pub/cet/issue/29516/316674>

Appendix A

Treatment schedule across 15 weeks

Week	Group A	Group B	Control Group
1	Pre-test survey + System Training Assignment 1		
2	WCF System Training (Count/Non-count nouns) Assignment 2		
3	Pre-test Assignment 3		
4	<i>to-verb</i> CF 2 → Assignment 4	<i>articles</i> CF 2 → Assignment 4	Feedback 2 → Assignment 4
5	<i>to-verb</i> CF 3 → Assignment 5	<i>articles</i> CF 3 → Assignment 5	Feedback 3 → Assignment 5
6	Mid-test <i>to-verb</i> CF 4 → Assignment 6	Mid-test <i>articles</i> CF 4 → Assignment 6	Mid-test Feedback 4 → Assignment 6
7	<i>articles</i> CF 5 → Assignment 7	<i>to-verb</i> CF 5 → Assignment 7	Feedback 5 → Assignment 7
8	<i>articles</i> CF 6 → Assignment 8	<i>to-verb</i> CF 6 → Assignment 8	Feedback 6 → Assignment 8
9	Post-test <i>articles</i> CF 7 → Assignment 9	Post-test <i>to-verb</i> CF 7 → Assignment 9	Post-test Feedback 7 → Assignment 9
10	Feedback 8 → Assignment 10		
11	Feedback 9 → Assignment 11		
12	Delayed-post-test Feedback 10 → Assignment 12		
13	Feedback 11 → Assignment 13		
14	Feedback 12		
15	Post-test Survey Feedback 13		

Appendix B

Pre-test survey results

Table B1. *Teacher's role*

Question	Strongly Agree	Moderate Agree	Mild Agree	Mild Disagree	Moderate Disagree	Strongly Disagree
27* Teachers should rewrite grammar errors	16.7%	46.7%	33.3%	3.3%	0.0%	0.0%
28 Teachers should explain grammar errors	26.7%	55.0%	18.3%	0.0%	0.0%	0.0%
29 Teachers should rewrite structural errors	13.6%	55.9%	27.1%	3.4%	0.0%	0.0%
30 Teachers should explain structural errors	22.0%	49.2%	23.7%	5.1%	0.0%	0.0%
31 Teachers should always leave positive comments	5.1%	32.2%	30.5%	25.4%	3.4%	3.4%

* $p < .1$

Table B2. *Detailed response to Question 27*

Question 27: Teachers should rewrite grammar errors						
	Strongly Agree	Moderate Agree	Mild Agree	Mild Disagree	Moderate Disagree	Strongly Disagree
Group A	11.76%	47.06%	41.18%	0.00%	0.00%	0.00%
Group B	37.50%	43.75%	12.50%	6.25%	0.00%	0.00%
Control	7.41%	48.15%	40.74%	3.70%	0.00%	0.00%

Table B3. *Student's role*

Question	Strongly Agree	Moderate Agree	Mild Agree	Mild Disagree	Moderate Disagree	Strongly Disagree
32 I always review teacher feedback.	35.6%	39.0%	18.6%	3.4%	3.4%	0.0%
33 I always review corrections provided by my teacher.	33.9%	47.5%	11.9%	5.1%	1.7%	0.0%
34 I review grammar topics I am weak at.	6.8%	32.2%	45.8%	13.6%	1.7%	0.0%
35 I review structure topics I am weak at.	6.8%	32.2%	44.1%	13.6%	3.4%	0.0%

Appendix C

Post-test survey results

Table C1. Moodle

	Group A (<i>n</i> = 16)	Group B (<i>n</i> = 13)	Control (<i>n</i> = 26)
1. Did you understand how to view your teacher feedback for graded assignments?			
Yes	100%	100%	92%
No	0%	0%	8%
2. Did you have any problems viewing the feedback?			
Yes	0%	8%	0%
No	100%	92%	100%
3. How often did you view your teacher feedback?			
Every time	19%	54%	12%
Almost every time	44%	31%	46%
Over half	31%	0%	31%
Less than half	0%	8%	12%
Almost never	6%	8%	0%
Never	0%	0%	0%
4. Did you understand how to view the grammar pages?			
Yes	100%	85%	
No	0%	15%	
5. Did you have any problems viewing the grammar pages?			
Yes	0%	0%	
No	100%	100%	
6. How often did you click on teacher grammar corrections to view the grammar pages?			
Every time	13%	15%	
Almost every time	19%	38%	
Over half	56%	23%	
Less than half	6%	8%	
Almost never	6%	15%	
Never	0%	0%	

Table C2. Teacher Feedback

	Group A (n = 16)	Group B (n = 13)	Control (n = 26)
8. Did you think the grammar correction of “articles” was helpful to you?			
Yes	100%	92%	
No	0%	8%	
9. Did you think the grammar pages on “articles” were helpful to you?			
Yes	100%	92%	
No	0%	8%	
10. Did you think the grammar correction of “to-infinitives” was helpful to you?			
Yes	100%	92%	
No	0%	8%	
11. Did you think the grammar pages on “to-infinitives” were helpful to you?			
Yes	100%	92%	
No	0%	8%	
12. Did you find your teachers comments helpful?			
Yes	100%	92%	100%
No	0%	8%	0%
13. Please select ONE statement that BEST describes how you feel about your English grammar.			
My English grammar problems are very serious and really hurt my writing.	44%	23%	54%
Although I don't know much about English grammar, it's not a serious problem for me.	38%	31%	23%
English grammar is not really a serious issue for me. Other writing issues are more important.	6%	23%	15%
I'm not really sure whether English grammar is a problem for me writing.	13%	23%	8%
14. In your opinion, what is the best way for your teacher to give your feedback about grammar errors in your writing? Please select ONE statement only.			
Don't correct my grammar. Let me try to correct my errors myself.	0%	0%	0%
Only correct the most serious errors.	0%	8%	4%
Circle or underline my errors, but don't correct them for me.	0%	0%	12%
Circle my errors and tell me what type of errors they are (verb tense, word choice, etc.).	81%	77%	54%
Correct all of my errors for me.	19%	15%	31%

¹ Regarding the focus of correction, nomenclature for experimental groups have been used rather loosely across the literature. WCF targeting multiple specific forms has been referred to as both *unfocussed* (Sheen et al., 2009) and *focused* (Shintani et al., 2013). Furthermore, *unfocussed* has at times also referred to groups that have received partial targeted correction (Ellis, et al., 2008). However, there is a clear contrast between these and Chandler (2003), who used a form of *unfocussed* WCF which is now often termed *comprehensive* feedback (i.e., where all errors come under scrutiny). For clarity, the author considers *unfocussed* to be synonymous with *comprehensive*.