

Article

Sustaining University English as a Foreign Language Learners' Writing Performance through Provision of Comprehensive Written Corrective Feedback

Xiaolong Cheng ¹ and Lawrence Jun Zhang ^{2,*} 

¹ School of Foreign Languages, Hubei University of Technology, Wuhan 430068, China; xl.cheng@auckland.ac.nz

² Faculty of Education and Social Work, The University of Auckland, Auckland 1023, New Zealand

* Correspondence: lj.zhang@auckland.ac.nz

Abstract: Writing is regarded as a crucial skill in English language curricula at the secondary and tertiary levels in the Chinese education system. Currently, Chinese teachers of English as a foreign language (EFL) often adopt a product approach to teaching EFL writing, in which they emphasize the quality of their students' written products and show little concern with the writing process. To help L2 learners achieve sustainable development of their writing proficiency, teachers employ a comprehensive approach to correct their students' language errors as a common practice. However, empirical studies regarding its efficacy on different dimensions of L2 writing are insufficient. This study intended to fill this lacuna in a Chinese EFL context, which investigated the effects of sustained comprehensive written corrective feedback (WCF) on accuracy, complexity, fluency, and content and organization quality of EFL students' writing. Quasi-experimental in design, it involved a comparison group and a treatment group receiving four sessions of direct comprehensive WCF. Results show that such WCF contributed to writing accuracy and fluency over time. Our textual analysis further reveals that it particularly benefited students' grammatical accuracy, reducing some rule-based grammatical error types. However, it showed limited effects on complexity, content, or organization of students' writing. Interestingly, the comparison group did not improve any dimensions of their writing. Possible implications are also discussed.

Keywords: sustainable writing development; comprehensive written corrective feedback; L2 writing performance; Trade-off Hypothesis



Citation: Cheng, X.; Zhang, L.J. Sustaining University English as a Foreign Language Learners' Writing Performance through Provision of Comprehensive Written Corrective Feedback. *Sustainability* **2021**, *13*, 8192. <https://doi.org/10.3390/su13158192>

Academic Editor: Fuhui Tong

Received: 1 June 2021

Accepted: 15 July 2021

Published: 22 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As an indispensable practice in writing instruction to sustain L2 learners' writing performance, feedback is widely utilized by L2 writing teachers to inform students of their writing problems and weaknesses so that students can improve their writing performance in both local (language) and global (content and organization) aspects [1–3]. Therefore, what invariably happens in L2 writing classrooms is that teachers spare no effort to provide students with feedback on a variety of errors, particularly errors in language use (i.e., written corrective feedback, or WCF for short) [4–6].

In the existing literature, there is a spirited discussion among researchers on the efficacy of WCF, which was triggered by Truscott [7]. After the synthesis of some early studies, he vehemently repudiated the practice of WCF, claiming that it was not only ineffective but also harmful for L2 writing [8,9]. However, a number of researchers refuted his claim, documenting that WCF can contribute to L2 writing accuracy in both revised and new texts [10–15]. Such a controversy reflects the complexity of WCF in the sphere of L2 writing. More significantly, what adds the complexity to this issue is the extent to which WCF should be provided in order to maximize its effectiveness. Some researchers have advocated that a focused (selective) feedback approach is supposed to be adopted.

That is, teachers only focus on one or a few pre-selected types of errors and leave others uncorrected [14,16–18]. Other scholars, in contrast, have cast doubts on the practicability of this approach in authentic L2 writing contexts [19–23]. From their perspective, unfocused (comprehensive) written feedback (i.e., the correction of a wide array of linguistic errors) should deserve a place in writing instruction. Apparently, a controversial issue among WCF researchers still remains, concerning the scope of written feedback that should be provided by teachers.

Despite the debate over focused and comprehensive WCF, the latter has become a ubiquitous practice in L2 writing teaching [5,24–26]. While recent years have witnessed the proliferation of studies on focused WCF, studies that take comprehensive WCF as the research focus are limited [13,25,27,28]. Considering that studies on comprehensive WCF are relatively scarce and it has great relevance to real L2 writing classrooms, we report a quasi-experimental study, which concentrated on and tracked what effects of comprehensive WCF had on Chinese EFL learners' writing performance. In addition, different from the most existing WCF studies that have used a single or limited measures to evaluate students' writing production, our study employed multiple measures (i.e., accuracy, linguistic complexity, fluency, content and organization quality) due to the possible trade-off among them (see details in Section 2.3), which could advance the current knowledge in this area. Pedagogically, our study would enable WCF practitioners to reflect on their current WCF practices and optimize their WCF practices or initiate some pedagogical innovations to help students make progress in different dimensions of L2 writing, which is regarded as a sustainable goal in L2 writing pedagogy.

Since this study was conducted in the Chinese tertiary EFL context, we offer some important background information on EFL education in Chinese universities. At the tertiary level, EFL teaching in mainland China follows the national curriculum requirements [6,29], which are mandated by the Ministry of Education of China [3]. Upon admission to universities, all students are required to learn English as a foreign language for at least two academic years. In general, their teachers teach them English with a grammar translation approach, whereby linguistic accuracy is emphasized [30]. In addition, Chinese tertiary EFL learners also have to sit for some high-stakes examinations to demonstrate that they are qualified EFL learners. These examinations include the College English Test, Bands 4 and 6 (CET-4 and CET-6), and the Test for English Majors, Bands 4 and 8 (TEM-4 and TEM-8). Overall, English examinations are of great importance for tertiary EFL teaching and learning in China. In other words, EFL education in China is, to some degree, examination-oriented [31].

2. Literature Review

2.1. Is WCF Effective in L2 Writing?

As a typical representative calling the value of WCF into question, Truscott [7–9] published a series of articles to argue that WCF does not play a role in L2 writing. He claimed that WCF not only does not show any effects on writing accuracy but also may produce simplified and short writing, which makes linguistic complexity and fluency suffer. Thus, teachers should abandon this pedagogical practice in writing instruction. However, his claim was rejected by a great many WCF researchers, who have lent a plethora of empirical evidence to demonstrate that WCF is of great value in L2 writing. Specifically, not only can WCF contribute to students' text revision [13,32,33], but also has positive impacts on new pieces of writing [34–36]. Favorable effects of WCF on new pieces of writing illustrate that its effects go beyond the revision stage. In the process of back-and-forth debate over its effectiveness, WCF has become a significant arena in L2 writing for robust investigations [25].

Scholars are in favor of the view that WCF still needs to be emphasized and employed in the learning and teaching of L2 writing [37–39]. In order to maximize the efficacy of WCF, researchers have paid much attention to the strategies to deliver WCF (direct or indirect WCF) as well as the scope of WCF to be offered (focused or unfocused WCF). Since

this study is concerned with the effects of comprehensive WCF, the studies relevant to focused and unfocused WCF are reviewed next.

2.2. *Focused or Unfocused WCF: Which One Is More Effective?*

In the extant WCF literature, a divergence exists between focused (selective) feedback and unfocused (comprehensive) feedback. The focused-unfocused dichotomy hinges on the “the comprehensiveness of correction” (p. 11) [40]. Currently, a great many scholars in WCF show preference for focused WCF over comprehensive WCF and they have presented theoretical as well as empirical justifications for their preference. Theoretically, targeting only a few linguistic structures is friendlier to L2 learners, especially those under-achieving ones, for it can avoid their cognitive overload and enable them to have more attentional resources to process new input [27,36]. Empirically, the researchers in this line have testified the effectiveness of focused WCF on L2 writing performance [10,14,35,41,42]. These studies revealed that the recipients of focused WCF outperformed those who did not receive WCF in terms of writing accuracy in the target structure(s), verifying the value and usefulness of focused WCF.

Although WCF researchers provided extensive empirical evidence to shed light on the effectiveness of focused WCF, some scholars voiced their apprehension towards this practice due to its lack of ecological validity [13,37,43,44]. The existing literature on focused WCF tends to target one or two linguistic features. From a practical perspective, L2 learners tend to commit a variety of linguistic errors in writing, and the ultimate goal of teachers’ providing WCF is to improve the overall writing accuracy rather than accuracy in specific linguistic categories [37,45]. Thus, targeting only one or a few linguistic errors is far from sufficient. To generate direct pedagogical implications, the research on WCF should better reflect what happens in authentic L2 writing classrooms [13,46]. In view of this, comprehensive WCF, which corresponds to the actual WCF practices in classroom settings, may deserve due attention in scholarship on writing [40,44,47].

Although several studies have touched upon comprehensive WCF, with a research design that had one treatment group receiving comprehensive WCF, to be compared against other treatment/comparison groups [11,18,48], they were not systematically executed. The reason is that the research focus of these studies was not whether comprehensive WCF affected the general accuracy in L2 learners’ writing, with which writing teachers are concerned. As aforementioned, the recent literature that exclusively focuses on the effects of comprehensive WCF is much limited in number when compared with studies on focused WCF, and consequently, inconsistent research findings are documented [13,20–22,49]. For instance, Truscott and Hsu [49] inquired into the effects of comprehensive WCF with 47 upper-intermediate ESL learners as participants. They found that comprehensive WCF had favorable effects on the general accuracy in revision, but it failed to transfer such effects to new pieces of writing. Similarly, Karim and Nassaji [13] found that comprehensive WCF enabled L2 learners to improve their accuracy in revised writing, but it did not yield significantly beneficial effects on general accuracy in new texts. Such results indicate that comprehensive WCF is just an effective editing tool. By contrast, Van Beuningen, De Jong, and Kuiken [21,22] examined comprehensive WCF in Dutch as a foreign language (DFL) settings. They reported that such a practice not only contributed to students’ accuracy in revised writing, but also enabled them to improve their performance in accuracy in new pieces of writing. Their results suggest that comprehensive WCF is both a valuable editing and learning instrument. The mixed results regarding the effects of comprehensive WCF on new texts may be attributed to various factors such as the complexity of WCF, research settings, participants’ linguistic proficiency, and the inconsistent error categories that WCF targeted [13,50].

2.3. *Trade-Off Hypothesis*

As a well-known theory in the realm of research on task complexity, Skehan’s [51,52] Trade-off Hypothesis, which aligns with the theories of working memory, informs the task

design and implementation, as well as makes predictions for the effects of task manipulation on L2 production [53]. This hypothesis, operationalized in the Limited Attention Model, postulates that L2 learners have limited attentional resources in task execution, so they can only attend to one aspect of language (i.e., complexity, accuracy, and fluency, CAF for short) at a given time in performing a task [51,52]. In this sense, it is likely that they improve one area of language production (CAF) at the expense of others. That is, there may exist a competition among the three features of L2 production.

From the perspective of the Trade-off Hypothesis, it could be predicted that WCF makes L2 learners pay much attention to the accuracy of linguistic forms and they probably have few extra attentional resources available to process linguistic complexity and increase linguistic fluency, which may hinder their development in these two dimensions. In other words, the improvement in accuracy rendered by WCF may compromise the linguistic complexity and fluency [15]. Such a prediction is in alignment with Truscott's [8] claim that in order to enhance accuracy, WCF recipients would shorten and simplify their writing. To date, a few WCF studies have assessed the accuracy along with complexity and/or fluency. Of the limited studies, this prediction was supported by Kepner [54] and Semke [55]. However, it was contradicted by others [20,22,56,57]. For example, Rahimi [57] found that students who were under the treatment of error correction improved the syntactic complexity, as well. Regarding fluency, it was revealed by Chandler's [56] study that WCF did not affect the writing fluency negatively. Instead, such a practice benefited its development.

Aside from the trade-off among CAF, Skehan [51,52] also posits that there is a competition between language and high-order dimensions in relation to L2 production. Like the above prediction on CAF, it can be hypothesized that WCF directs students' attention to language and results in the corresponding enhancement, which may induce little attention to high-order dimensions of writing (e.g., content and organization), imposing a harmful impact on these aspects. At present, not much is known about the effects of WCF on the content and organization of the text that students generate [58]. Two studies, to our knowledge, have attempted to address this issue, and the researchers found that WCF did not show any effects on the rhetoric appropriateness in writing although it contributed to the accuracy significantly [20,46]. Interestingly, the findings are not supportive of the trade-off effect between language and high-order dimensions.

In general, the current body of literature paints an incomplete picture about WCF effects [20,57,58]. Specifically, the prior studies mainly concentrate on whether such a methodology impacts accuracy but pay little attention to its effects on other aspects of writing. According to the Trade-off Hypothesis, the gains in accuracy probably come at the cost of other dimensions in relation to writing production. Thus, in order to contextualize the writing accuracy, other important indexes related to writing performance such as complexity, fluency, content, and organization need to be considered in WCF research [15,46].

To sum up, our review suggests that there is a need to further investigate comprehensive WCF. Firstly, in comparison with focused WCF, the recent studies focusing on comprehensive WCF are less common and have produced conflicting results concerning its effects on new pieces of writing [14,59]. Therefore, more studies are warranted to clarify the effects. Another concern is the length of WCF treatment. Most existing WCF studies tend to employ a one-shot intervention to explore its efficacy. Although such a design is relatively easy to implement, it is rather questionable [13,43]. Specifically, students need exposure to extensive WCF to enhance their writing proficiency [44], and the development of writing skills is not a linear but recursive process. In this sense, it would be unrealistic to expect L2 learners to improve their writing performance and maintain such improvements over a short period of time, let alone a one-off WCF [43]. Thus, multiple WCF treatment sessions are called for to promote the ecological value and have a better understanding of the effects of sustained WCF. Lastly, as noted earlier, the majority of prior studies adopt accuracy as the one single measure to assess WCF effects [60]. This practice in research

fails to tell the whole story, as its effects on other aspects of writing performance should be examined due to the potential trade-off.

With the above considerations in mind, this study set out to address the following research questions with an ameliorated research design:

- (1) What are the effects of comprehensive WCF on Chinese EFL learners' writing complexity, accuracy, and fluency?
- (2) What are the effects of comprehensive WCF on Chinese EFL learners' content and organization quality?

3. Methods

3.1. Context and Participants

This study was conducted at a medium-ranked university in mainland China. Two parallel intact classes of English major sophomores ($N = 72$) were recruited on the basis of a convenience sampling strategy [61]. At this university, these students were required to enroll in an *English Writing Course*, which was offered once a week (two 45-min class periods) in a 16-week semester. Such a course was designed to develop students' knowledge in English writing and foster their writing competence in some basic genres (i.e., narrative, argumentation, and exposition). The course for these two classes was taught by the same teacher, who earned her master's degree in applied linguistics and had eight years of EFL writing teaching experience. The two classes were randomly assigned to either a treatment group ($N = 36$) or a comparison group ($N = 36$). The two groups used the same university syllabus and attended the same writing activities both in and after class.

Before the intervention, all the participants completed a demographic questionnaire. In total, there were nine males (12.50%) and 63 females (87.50%) and their ages ranged from 19 to 21 with, on average, 10 years of English learning experience. In addition, all the participants learned English in mainland China and had no experience of studying abroad. In a word, the participants in the two groups were comparable in terms of age, educational background, and English learning experience. According to their examination scores in the last semester and the discussion with the course teacher, we came to a decision that these participants should be viewed as intermediate EFL learners.

3.2. Data Collection

This quasi-experimental study was over a span of nine weeks, consisting of three testing sessions and four rounds of treatment sessions. The detailed procedures of how the testing and treatment sessions were conducted are presented below (see Table 1).

Table 1. Procedures of data collection.

Week		Treatment Group	Comparison Group
1	Day 1	Pretest (text 1)	Pretest (text 1)
	Day 3	WCF + rewriting text 1	Rewriting text 1
2	Day 1	Text 2	Text 2
	Day 3	WCF + rewriting text 2	Rewriting text 2
3	Day 1	Text 3	Text 3
	Day 3	WCF + rewriting text 3	Rewriting text 3
4	Day 1	Text 4	Text 4
	Day 3	WCF + rewriting text 4	Rewriting text 4
5		Posttest	Posttest
9		Delayed posttest	Delayed posttest

3.2.1. Writing Tests

In our study, we chose argumentation as the genre for the writing tests. Argumentative writing is recognized as a reliable and popular instrument to evaluate L2 learners' writing proficiency in academic contexts [31]. As such, Chinese tertiary EFL learners tend to be required to complete argumentative writing in various well-established English proficiency

examinations, including IELTS, TOEFL, College English Test Band 4/6 (CET-4/6), and Tests for English Majors Band 4/8 (TEM-4/8).

The participants in the two groups were invited to take a pretest, a posttest, and a delayed posttest, which were given prior to as well as immediately after the treatment and three weeks after the posttest. These tests were used to examine the effects of comprehensive WCF on the different aspects of Chinese EFL learners' writing performance. Three different topics for the three tests were selected from the past TEM-4 test battery (see Appendix A). Such a decision was made based on two rationales. Firstly, teachers' instructional practices are always examination-driven in Chinese tertiary EFL writing classrooms [3], so their students tend to write arguments based on the writing topics from the past TEM-4 papers. More importantly, since TEM-4 is a large-scale standardized test in mainland China, its writing topics are of high reliability and validity. To be specific, they are drawn from general education and students' everyday life experiences, so they are considered familiar and fair to each student [4], which ensures the difficulty of such topics to be largely consistent.

All the participants were required to complete each of the writing tasks within 40 min and the length of writing was expected to be no less than 200 words, as stipulated by TEM-4. During the testing sessions, they did not have access to any external resources such as dictionaries and textbooks.

3.2.2. Intervention

In the treatment group, participants received four rounds of WCF targeting all errors (e.g., articles, verbs, pronouns, singular/plural forms, part of speech, sentence structures, and word choices, among others). When conducting research on WCF, we should take the type(s) of WCF into consideration. In general, there are two types of WCF: direct WCF and indirect WCF. Specifically, the former refers to teachers' providing students with correct forms, while indirect WCF is realized by only indicating errors without offering students correct forms. When it comes to our study, we provided students with direct feedback (see following examples). This practice was rationalized by several reasons. First, in comparison with indirect WCF, direct WCF provides students with input, which enables them to avoid misunderstanding and confusion while coping with WCF, internalize the correct forms immediately, and have explicit information to test the hypotheses that they have made on target language [36,40,56]. Second, considering that the participants in our study were intermediate EFL learners instead of advanced ones, direct WCF was probably more appropriate for them. As suggested by scholars [62,63], indirect WCF is better suited to students with high language proficiency. Finally, whereas the relative effectiveness of direct and indirect WCF is inconclusive, Kang and Han's [50] meta-analysis showed that direct WCF is more effective than indirect WCF with regard to facilitating linguistic accuracy ($g = 0.60$ vs. 0.33).

Example 1. *It can make an efficient world which we live . . . (in which).*

Example 2. *Otherwise, intelligent machines will arise our laziness . . . (arouse).*

Example 3. *A intelligent machine can only take place the elementary work . . . (An).*

After each round of feedback provision, the students received the first draft of their essays with WCF. The students were given 10 min to study the provided WCF. After they had looked over their feedback, their writing samples were collected and they were allocated another 30 min to rewrite without having access to WCF. We followed many previous studies which involved direct WCF [13,47] to administer such a feedback procedure and its purpose was to avoid students' shallow processing of WCF (i.e., just copying corrections) because of direct WCF and make students engage with WCF more deeply [64]. In the process of rewriting, they were not provided with any external assistance.

The participants in the comparison group did not receive any feedback and they were just required to complete their writing based on the same given writing prompts. After finishing the writing tasks, they were encouraged to self-correct their written scripts. In order to avoid disadvantaging the participants in the comparison group, we provided them with the WCF after the intervention.

3.3. Data Analysis

In this intervention, the dependent variables included syntactic and lexical complexity, accuracy and fluency; content and organization quality, which were assessed by different measures. To explore the effects of comprehensive WCF on syntactic complexity, two measures were adopted: mean length of T-units and ratio of clauses per T-unit, two widely-used measures for overall syntactic complexity. For lexical complexity, we also used two measures. One was lexical density, i.e., the proportion of content words (e.g., verbs, nouns, adjectives and adverbs) in the total number of words. The other was mean segmental (50 words as a segment) type-token ratio. Syntactic and lexical complexity were processed by the L2 Syntactic Complexity Analyzer (L2SCA) and Lexical Complexity Analyzer (LCA), two web-based programs [65,66]. As for fluency, it was appraised by the total number of words students produced within 40 min.

Accuracy was evaluated by the number of errors per 100 words. In this study, errors were defined as any grammatical, lexical, and morphological inaccuracies in students' writing [67]. However, errors in relation to spelling and punctuation were not included, which was based on two considerations. For one thing, such an exclusion was to avoid the possible over-estimation of errors due to the illegible handwriting; for another, our browsing of participants' writing showed that they made few errors in spelling and punctuation. In this study, we adopted Geng's [68] guidelines to code and count the errors. To have a nuanced understanding about the effects of comprehensive WCF on writing accuracy, we further analyzed which error types were amenable to comprehensive WCF, as recommended by the prior studies [27,59,63]. Initially, we conducted a broad classification of the errors (i.e., grammatical vs. lexical errors) in the pretest, posttest, and delayed posttest according to Van Beuningen et al.'s [22] coding scheme. Next, we implemented a narrow coding to examine the specific error types (e.g., article, subject-verb agreement, tense, and plural/singular form, among others). The narrow coding was partly based on Chandler [56] and partly data-driven.

The content and organization quality were scored based on Jacobs et al.'s [69] ESL Writing Profile, a well-established and widely-used L2 writing rubric, by which we could determine whether WCF would impact the high-order dimensions of writing. In this writing rating scheme, the scores in the content and organization ranged from 13 to 30 points and from 7 to 20 points, respectively.

In order to maintain the reliability of coding linguistic errors, a co-coder who was a Chinese tertiary EFL teacher with a master's degree in applied linguistics was invited. After the first author completed the coding work, 32 students' texts (approximately 15% of all the writing samples) were selected randomly and re-coded by her. The inter-coder reliability was calculated by intraclass correlation coefficient (ICC). The result showed that the inter-coder reliability for coding was 0.83, which exceeded the acceptable level of 0.7 [70]. The same approach was applied to ascertain the inter-rater reliability for content and organization quality. The inter-rater reliability for these two aspects turned out to be 0.90 and 0.89, respectively.

Prior to inferential statistical analyses, the normality of each dependent variable was firstly scrutinized. In this study, the data were assumed to be normal distributed when the z-scores of skewness and kurtosis are less than 1.96 [71]. Since the data of our study met the normal distribution requirement, ANOVAs and *t*-tests were employed.

As for the calculation of the effect size, Cohen's *d* was used for *t*-tests and partial η^2 for ANOVAs. According to Cohen's [72] criteria, *d* values of 0.20, 0.50, and 0.80 and partial η^2 values of 0.01, 0.06, and 0.14 were considered to be small, medium, and large, respectively.

4. Results

Table 2 lists the descriptive data of means and standard deviations for different measures between the two groups across three tests. In order to guarantee the baseline conditions between the treatment group and the comparison group at the outset of the intervention, independent samples *t*-tests were run (see Table 3). The tests showed that there were no significant differences with respect to the various indexes at the time of the pretest: Mean length of T-units, $p = 0.333$; ratio of clauses per T-unit, $p = 0.103$; lexical density, $p = 0.997$; mean segmental type-token ratio, $p = 0.088$; errors per 100 words, $p = 0.970$; fluency, $p = 0.475$; content, $p = 0.721$; organization, $p = 0.693$.

Table 2. Descriptive statistics for different measures over time.

Measures	Group	Pretest		Posttest		Delayed Posttest	
		Mean	SD	Mean	SD	Mean	SD
MLT	TG	14.58	2.29	14.17	2.35	15.56	2.25
	CG	13.92	2.44	14.00	2.04	14.26	1.77
RCT	TG	1.61	0.19	1.52	0.23	1.64	0.29
	CG	1.50	0.28	1.51	0.23	1.59	0.18
LD	TG	0.55	0.030	0.56	0.029	0.54	0.031
	CG	0.55	0.040	0.55	0.036	0.52	0.027
MSTTR	TG	0.79	0.039	0.79	0.033	0.78	0.035
	CG	0.77	0.045	0.78	0.035	0.76	0.037
EP100W	TG	7.96	2.22	5.27	1.65	4.79	1.82
	CG	7.98	2.63	9.53	2.77	8.41	2.35
Fluency	TG	213.24	32.47	237.68	39.41	240.72	31.36
	CG	220.25	35.68	218.71	39.68	217.75	32.73
Content	TG	18.04	1.62	18.72	2.05	18.68	2.04
	CG	17.76	1.60	18.75	1.75	18.93	1.27
Organization	TG	13.36	1.35	13.64	1.50	13.88	1.51
	CG	13.21	1.32	13.88	1.23	13.67	1.37

Note: TG = treatment group; CG = comparison group; MLT = mean length of T-units; RCT = ratio of clauses per T-unit; LD = lexical density; MSTTR = mean segmental type-token ratio; EP100W = the number of errors per 100 words; SD = standard deviation.

Table 3. Between-subjects comparisons for different measures across tests.

Measures	Pretest		Posttest		Delayed Posttest	
	<i>t</i>	<i>p</i>	<i>t</i>	<i>p</i>	<i>t</i>	<i>p</i>
MLT	0.977	0.333	0.264	0.793	2.232	0.030 *
RCT	1.622	0.103	0.216	0.830	0.703	0.486
LD	0.003	0.997	0.633	0.530	2.134	0.038 *
MSTTR	1.740	0.088	1.067	0.291	2.282	0.027 *
EP100W	−0.038	0.970	−6.50	0.000 **	−6.05	0.000 **
Fluency	−0.720	0.475	1.679	0.100	2.509	0.016 *
Content	0.359	0.721	−0.055	0.956	1.063	0.294
Organization	0.398	0.693	−0.600	0.552	0.517	0.608

Note: MLT = mean length of T-units; RCT = ratio of clauses per T-unit; LD = lexical density; MSTTR = mean segmental type-token ratio; EP100W = the number of errors per 100 words. * $p < 0.05$; ** $p < 0.001$.

4.1. Effects on Mean Length of T-Units (MLT)

As seen in Table 3, there are no significant differences between the two groups in terms of MLT in the pretest ($t = 0.977$, $p = 0.333$) and the posttest ($t = 0.264$, $p = 0.793$), but the difference is notable in the delayed posttest ($p = 0.030$) with a medium effect size ($d = 0.64$).

One-way repeated measures ANOVAs showed that the mean scores of MLT developed differently across time in the treatment group ($F(2, 70) = 4.134$, $p = 0.022$, partial $\eta^2 = 0.147$), but not in the comparison group ($F(2, 70) = 0.267$, $p = 0.767$). To further investigate the within-subjects differences in the treatment group, a series of paired samples *t*-tests were

employed and a Bonferroni correction was used for the three comparisons ($p = 0.017$). The paired samples t -tests showed that the treatment group displayed significantly better results in MLT from the posttest to the delayed posttest ($p = 0.002$, $d = 0.72$), but no significant improvement was observed from the pretest to the posttest, nor from the pretest to the delayed posttest.

4.2. Effects on Ratio of Clauses per T-Unit (RCT)

Table 3 reveals that the treatment group and the comparison group had similar performance in terms of RCT in the pretest ($t = 1.662$, $p = 0.103$). In the same vein, no statistically significant differences were found in the posttest and the delayed posttest between the two groups ($t = 0.216$, $p = 0.830$; $t = 0.703$, $p = 0.486$, respectively). One-way repeated measures ANOVAs showed that the means of RCT did not vary significantly over time for the treatment group ($F(2, 70) = 2.576$, $p = 0.087$) or the comparison group ($F(2, 70) = 1.596$, $p = 0.218$).

4.3. Effects on Lexical Density (LD)

As shown in Table 3, the performance of the two groups regarding LD in the pretest and the posttest did not vary greatly ($t = 0.003$, $p = 0.997$; $t = 0.633$, $p = 0.530$, respectively). However, the treatment group outdid the comparison group in the delayed posttest ($p = 0.038$) with a medium effect size ($d = 0.61$). In terms of within-subjects differences, one-way repeated measures ANOVAs demonstrated no significant differences of LD in the treatment group over time ($F(2, 70) = 2.011$, $p = 0.145$), but occurred in the comparison group ($F(2, 70) = 6.184$, $p = 0.004$, partial $\eta^2 = 0.212$). Specifically, the mean scores of LD in the comparison group decreased significantly from the posttest to the delayed posttest ($p = 0.006$, $d = 0.623$) and from the pretest to the delayed posttest ($p = 0.006$, $d = 0.624$).

4.4. Effects on Mean Segmental Type-Token Ratio (MSTTR)

The between-subjects comparisons revealed that the mean scores of MSTTR between the two groups were similar in the pretest and the posttest ($t = 1.740$, $p = 0.088$; $t = 1.067$, $p = 0.291$, respectively) (see Table 3). However, students in the treatment group outperformed their comparison group peers with regard to MSTTR in the delayed posttest ($p = 0.027$, $d = 0.65$).

Regarding the within-subjects differences, one-way repeated measures ANOVAs showed that neither the treatment group nor the comparison group changed the means of MSTTR significantly across tests ($F(2, 70) = 1.069$, $p = 0.352$; $F(2, 70) = 2.887$, $p = 0.066$, respectively).

4.5. Effects on Number of Errors per 100 Words (EP100W)

Table 3 shows that the treatment and comparison groups had similar performance regarding EP100W in the pretest ($t = -0.038$, $p = 0.970$). Two-way repeated measures ANOVAs found that there were significant main effects for time and group ($F(2, 140) = 5.787$, $p = 0.006$, partial $\eta^2 = 0.110$; $F(1, 70) = 34.292$, $p = 0.000$, partial $\eta^2 = 0.422$, respectively). The time-group interaction effect was significant as well ($F(2, 140) = 15.926$, $p = 0.000$, partial $\eta^2 = 0.253$). This indicated that the treatment group and the comparison group differed greatly in their performance in this measure over time. To further explore the between-subjects differences, another two independent samples t -tests were conducted, as shown in Table 3. Results showed that the two groups varied greatly with regard to EP100W in the posttest ($p = 0.000$) and the delayed posttest ($p = 0.000$) with large effect sizes ($d = 1.86$ and $d = 1.73$, respectively). This suggested that the treatment group had much fewer errors than the comparison group in both the posttest and the delayed posttest. To establish whether the changes were significant in each group over time, one-way repeated measures ANOVAs were computed. Results indicated that there were significant changes across tests in the treatment group ($F(2, 70) = 30.069$, $p = 0.000$, partial $\eta^2 = 0.556$), but no significant changes were observed in the comparison group over time ($F(2, 70) = 2.707$,

$p = 0.097$). The post-hoc analyses revealed that the intervention helped the students in the treatment group reduce the number of errors greatly from the pretest to the posttest ($p = 0.000$, $d = 1.22$) and from the pretest to the delayed posttest ($p = 0.000$, $d = 1.37$).

Apart from the examination of the effects on general writing accuracy, we explored an important question in L2 writing research and instruction: Which error types benefit from comprehensive WCF? In terms of the broad categories (grammatical vs. lexical errors), one-way repeated measures ANOVAs showed that such WCF had few effects on Chinese EFL learners' lexical accuracy ($F(2, 70) = 1.732$, $p = 0.188$), but it facilitated their grammatical accuracy significantly across the three tests (337 vs. 239 vs. 223; $F(2, 70) = 41.855$, $p = 0.000$, partial $\eta^2 = 0.636$). Regarding the specific error types, as Table 4 illustrates, direct comprehensive WCF prompted students to reduce the errors in relation to verb forms, articles, singular/plural forms, and run-on sentences from the pretest to the immediate posttest. More encouragingly, they maintained the positive effects on these error types in the delayed posttest. Overall, these errors were ruled-based grammatical features.

Table 4. Frequencies and percentages for error types across tests.

	Pretest	Posttest	Delayed Posttest
<i>Grammar</i>			
Verb forms	88(26.11%)	61(25.52%)	54(24.22%)
Articles	62(18.40%)	41(17.15%)	30(13.45%)
Singular/plural forms	45(13.35%)	28(11.72%)	19(8.52%)
Agreement	48(14.24%)	33(13.80%)	21(9.42%)
Run-on sentences	31(9.20%)	19(7.95%)	12(5.38%)
Others	63(18.69%)	57(23.85%)	87(39.01%)
Total grammatical errors	337(100%)	239(100%)	223(100%)
<i>Lexis</i>			
Wrong words and word choice	78(100%)	73(100%)	65(100%)
Total errors	415(100%)	312(100%)	288(100%)

4.6. Effects on Fluency

As seen in Table 3, the treatment group and the comparison group wrote a similar number of words in the pretest ($t = -0.720$, $p = 0.475$) and the posttest ($t = 1.679$, $p = 0.100$). However, their mean difference was great in the delayed posttest ($p = 0.016$, $d = 0.72$), suggesting that the treatment group excelled the comparison group in fluency in the delayed posttest.

Results of one-way repeated measures ANOVAs indicated that the variations in fluency were significant in the treatment group across tests ($F(2, 70) = 6.930$, $p = 0.002$, partial $\eta^2 = 0.224$), but the significant variations did not present themselves in the comparison group ($F(2, 70) = 0.037$, $p = 0.923$). The paired samples t -tests revealed that the intervention enabled the treatment group to improve their fluency in language production over time (pretest vs. posttest, $p = 0.008$, $d = 0.57$; pretest vs. delayed posttest, $p = 0.001$, $d = 0.79$).

4.7. Effects on Content Quality

Our statistical analyses (see Table 3) led us to the findings that there were no significant between-subjects differences in the content in the pretest, the posttest, and the delayed posttest ($t = 0.359$, $p = 0.721$; $t = -0.055$, $p = 0.956$; $t = 1.063$, $p = 0.294$, respectively). Likewise, one-way repeated measures ANOVAs demonstrated no significant within-subjects differences in the treatment group ($F(2, 70) = 2.277$, $p = 0.114$) and the comparison group ($F(2, 70) = 2.574$, $p = 0.087$).

4.8. Effects on Organization Quality

The between-subjects comparisons presented in Table 3 illustrated that the two groups produced the virtually identical organization quality in each test ($t = 0.398, p = 0.693$; $t = -0.600, p = 0.552$; $t = 0.517, p = 0.608$, respectively).

Regarding the within-subjects differences, one-way repeated measures ANOVAs did not show any significant mean differences in organization quality over time in the treatment group ($F(2, 70) = 2.105, p = 0.133$) or the comparison group ($F(2, 70) = 1.841, p = 0.170$).

5. Discussion

5.1. Effects on Accuracy

According to the results, offering comprehensive WCF not only prompted Chinese EFL learners to improve their general writing accuracy in the immediate posttest, but also retain the gains in the delayed posttest. This refutes Truscott [7,8], whose argument is that the practice of error correction does not benefit L2 writing and teachers should refrain from it. Our results regarding the beneficial effects of comprehensive WCF can also be observed in the previous studies [21,22,47]. Such favorable effects on accuracy in our study are not surprising, which may be attributable to several potential factors. Firstly, the comprehensive WCF in our study was delivered directly. As many researchers claimed [11,35,40], direct WCF provides students with explicit information, which helps them avoid misunderstanding and confusion while attending to WCF, and internalize the correct forms instantly. Due to its explicitness and immediacy, direct WCF consumes less cognitive resources, and it is considered less cognitively demanding compared to indirect WCF. Moreover, such facilitative effects may be associated with the feedback procedure. As noted previously, the participants in our research studied direct comprehensive WCF, but they did not have access to it while revising their writing. This practice probably encouraged students to reflect on and process WCF more profoundly than just copying teachers' direct corrections. With such a practice, what these students learned from comprehensive WCF could be strengthened. In addition, the multiple rounds of WCF treatment sessions possibly account for the results. Specifically, our study provided the students with four rounds of comprehensive WCF treatment, which probably enabled them to notice the errors that they were prone to making. Accordingly, they could avoid these errors in the follow-up writing tasks. However, our results run counter to Truscott and Hsu's [49] study, in which comprehensive WCF did not enhance students' overall linguistic accuracy in new pieces of writing. The mixed results may be ascribed to some variations between their study and our study in methodology such as the number of WCF treatments (one-off treatment vs. multiple rounds of treatments) and writing genres (narratives vs. argumentative essays).

Our investigation extends the previous studies on the usefulness of focused WCF [10,11,18,35,36,41]. Although studies in this line have confirmed the effectiveness of WCF focusing on one or a few structures (e.g., English article system or past tense), little is known concerning whether such a feedback practice could achieve any improvement in overall accuracy [27,62]. The positive effects of comprehensive WCF on accuracy contradict Sheen's [36] and Sheen, Wright, and Moldawa's [18] proposition that unfocused WCF has little learning potential, in that such WCF makes L2 learners so cognitively overwhelmed that they cannot process it effectively.

In addition to general writing accuracy, we examined the error types amenable to comprehensive WCF. Our study revealed that it contributed to Chinese EFL learners' reduction in grammatical errors significantly from the pretest to the delayed posttest, whereas lexical accuracy appeared to not benefit from such a practice. Our result contradicts Truscott's [8] argument that WCF is only effective for non-grammatical errors. Instead, it aligns with what Van Beuningen et al. [22] have found, and this result is quite understandable. Generally speaking, grammatical errors tend to be classified into treatable errors, which can be corrected by certain patterns or rules, while lexical errors into untreatable errors, which are idiosyncratic in nature and cannot be explained by rules [63]. As a result, it is challenging and difficult for students to reduce lexical errors, make progress in lexical accuracy,

and maintain such progress [63]. More importantly, our study further revealed that such WCF helped Chinese EFL learners decrease their errors related to verb forms, articles, singular/plural forms, agreement, and run-on sentences over time. This is an important contribution made by our study, which advances the prior literature in this line [13,21,22]. These previous studies only focus on the effects of comprehensive WCF on broad error categories (i.e., grammar vs. nongrammar) and does not take the specific error types into consideration.

As for the participants in the comparison group, those who just rewrote their compositions without being offered any WCF were not able to improve writing accuracy in the immediate posttest or the delayed posttest. This result does not concur with Rahimi's [57] finding that non-WCF receivers could improve accuracy over time. Two plausible reasons may account for the disparity. First, her study lasted four months, which was much longer than this study. Second, in her study, the students in the comparison group still received general feedback comments on grammar, which may arouse their attention to their grammatical problems, while this study did not provide the comparison group with any feedback. The comparison group's failure to improve accuracy supports the assertion that the limited and short-term writing practices without feedback cannot contribute to EFL learners' accuracy in writing [73]. In other words, the sheer engagement in writing practices without extra scaffolding is futile, which parallels the recommendation that L2 learners' writing practices entail external support such as teacher feedback or writing instruction [62].

5.2. Effects on Syntactic Complexity

As aforementioned, two measures were employed to gauge the syntactic complexity. As for mean length of T-units, although the treatment group outperformed the comparison group in this index at the time of delayed posttest with a medium effect size, the two groups' mean scores in mean length of T-units remained unaffected from the pretest to the posttest as well as from the pretest to the delayed posttest. Regarding the ratio of clauses per T-units, it failed to vary significantly across groups or tests. This suggests that even though comprehensive WCF generated beneficial effects on accuracy, it appeared to not impact syntactic complexity adversely, regardless of how it was measured. This rebuts Truscott's [7,8] proposal that focusing on accuracy would compromise syntactic complexity, for learners may produce simplified writing to avoid errors.

From a theoretical perspective and in light of the research results regarding accuracy and syntactic complexity, this study does not agree with the Trade-off Hypothesis posited by Skehan [51,52], in which a competition between accuracy and complexity is expected. However, we found that students did not yield simplified writing alongside the improvement in accuracy. This reveals that there seems to be no trade-off between accuracy and syntactic complexity. In contrast, the results regarding the effects of accuracy and syntactic complexity in our study seem to be in favor of Robinson's Cognition Hypothesis [74,75], which assumes that complexity and accuracy are not in a trade-off relationship since these two dimensions, as two aspects of L2 learners' output, are considered to be connected closely [40,67,76].

Empirically, the finding that comprehensive WCF did not affect students' syntactic complexity negatively or positively coincides with the prior studies [19,22,46]. All these studies found that with the comprehensive WCF treatment, participants' syntactic complexity remained untouched. Two possible reasons might explain the unchanged syntactic complexity in our study. One is that the intervention was not long enough. This could be borne out by the evidence in Rahimi [57]. In her study, after an intervention over a long period of time, the participants receiving WCF treatment improved their syntactic complexity. The other reason may be attributed to the inadequacies of the two measures adopted to appraise students' syntactic complexity development in the present study. To date, a firm conclusion has not yet been drawn that these two measures are good indicators

of syntactic complexity development after a short period of time, notwithstanding their popularity in evaluating global syntactic complexity.

5.3. Effects on Lexical Complexity

Similar to the effects on syntactic complexity, no significant changes in lexical complexity were observed in the treatment group across tests even though it outperformed the comparison group in lexical density and mean segmental type-token ratio in the delayed posttest.

This finding corresponds to Hartshorn and Evans [46] and Van Beuningen, De Jong, and Kuiken [22], and is likely to be explained by two possible reasons. Firstly, our study explicitly required that students should complete writing tasks which were no less than 200 words within 40 min. In such an operationalization, they might consider lexical complexity comparatively peripheral and were likely to employ familiar words to complete the writing tasks, which resulted in the failure to improve lexical complexity. This indicates that they may sacrifice the lexical complexity for meeting the minimum requirements imposed by this study. More importantly, comprehensive WCF failed to enrich students' limited lexical knowledge. It focuses on and contributes to writing accuracy, but it may not play a role in enriching students' repertoire of vocabularies. Without a good command of lexical resources, it would be taxing for them to improve lexical complexity even if they feel there is a need to use more sophisticated or different types of words [77].

The ineffectiveness of WCF in improving lexical complexity can be viewed in an optimistic lens. This could be a good reason for L2 writing teachers to implement some pedagogical interventions, in which they need to include more sources of input alongside WCF provision. Doing so would afford students opportunities to expand their lexical resources and improve the writing accuracy simultaneously.

5.4. Effects on Fluency

Under the treatment of direct comprehensive WCF, students were able to write lengthier texts in the immediate posttest and the delayed posttest. Taking its effects on accuracy and fluency together, a conclusion can be drawn that this WCF contributed to L2 writers' writing fluency and accuracy concurrently. That is, the improvement of accuracy was not at the expense of fluency, which dismisses Skehan's [51,52] argument that there is a potential competition between accuracy and fluency.

The result that such WCF impacted fluency positively fits with what was reported by Chandler [56]. However, such a result differs from some previous studies, which reported that comprehensive WCF imposed few effects on students' fluency [19,20,46]. The inconsistent results are probably due to two potential factors. To begin with, compared with the participants in Hartshorn and Evans [46], participants in our study had a higher level of English proficiency as they were English major sophomores. Higher English proficiency endowed them with a greater potential to write longer texts. Moreover, this study required participants to complete the writing tasks with no less than 200 words within 40 min. To satisfy the requirement, students might try their best to write as much as possible. This differs from the research design in some previous studies [20,46]. In these studies, there was no requirement regarding the minimum number of words that students should accomplish.

5.5. Effects on Content and Organization Quality

As high-order dimensions of EFL writing production, both the content and the organization quality did not vary across tests in the treatment group. This suggests that comprehensive WCF did not show effects on them, which echoes the previous literature [20,46]. In this sense, it appears that such WCF exerts few effects on high-order dimensions of writing even though it can contribute to some linguistic output (i.e., accuracy and fluency), which responds to the claim that the improvement in linguistic areas does not pose a threat to the global level of L2 writing [53]. Such a finding does not lend support to the trade-off effects between language and high-order dimensions of L2 production [51,52]. In addition, this

study also did not find significant mean differences concerning content and organization quality in the comparison group over time. That the treatment and comparison groups did not make progress in content confutes Ashwell's [32] assumption that rewriting could help students produce better content quality.

The inefficacy of WCF on the content and the organization quality is not surprising. In general, after WCF provision, L2 learners engage themselves with correcting linguistic errors in that they prioritize such errors in revision [78,79]. With this flawed task revision schema, students are not very likely to identify their problems in global areas of writing. Moreover, this result may also be associated with the nature of global issues. Compared with linguistic errors, problems in content and organization are difficult to detect and solve by L2 learners independently, which require more of their cognitive resources [67,80]. Without external scaffolding, it is demanding for them to make progress in these aspects.

6. Conclusions

This quasi-experimental study examined the effects of comprehensive WCF on L2 writing performance. Adopting a quantitative approach, our research reveals its complex effects on Chinese EFL learners' writing production. Specifically, providing such WCF improved students' overall writing accuracy greatly in the immediate posttest and the favorable effects were sustained in the delayed posttest. Furthermore, while such WCF appeared to show no effects on lexical accuracy, it contributed to grammatical accuracy significantly. In terms of the specific error types, it benefited the rule-based grammatical errors such as errors related to verb forms, articles, and run-on sentences. Additionally, it had favorable effects in terms of increasing the number of words in writing in the short term and long term. However, this practice did not result in any significant changes in syntactic complexity, lexical density and diversity, or content and organization quality.

Unsurprisingly, this study was not free from limitations. Firstly, this study only focused on the effects of direct comprehensive WCF and did not take into account indirect comprehensive WCF, although such a decision was justified previously. However, given that the effects of WCF are mediated by the explicitness of WCF [40,50], we still have little knowledge about what effects indirect comprehensive WCF has on the different dimensions of L2 learners' writing. As such, further studies are warranted to examine this issue. Furthermore, the interval between the posttest and the delayed posttest was only three weeks. To better assess the delayed effects, future studies can extend the gap between the posttest and the delayed posttest, which could provide us with a deeper insight into comprehensive WCF effects over time. Finally, this study only documented the effects of comprehensive WCF on EFL learners' performance in argumentative writing. As Kang and Han [50] argued, WCF effects were influenced by task genres. Thus, future research needs to investigate the effects of comprehensive WCF on other genres.

Nonetheless, some contributions and implications can be derived from the present study. Although comprehensive WCF is often criticized for overwhelming L2 learners with multiple errors and imposing a heavy cognitive burden [5,10,18,36], our study reports that it benefits students' writing accuracy and fluency when offered directly. In this sense, direct feedback seems to reduce the cognitive load placed by comprehensive WCF. This suggests that to have a deep insight into the effectiveness of comprehensive WCF, feedback scope is not the only variable and the saliency of feedback (i.e., feedback explicitness) also plays an important role [47]. As claimed by many researchers [13,47,81], the explicitness of feedback matters when researchers examine WCF effectiveness. In addition, framed within the Trade-off Hypothesis, this study investigated the effects of comprehensive WCF on different dimensions of L2 writing. However, our results do not corroborate it. Specifically, we found that this practice improved L2 learners' accuracy and fluency simultaneously without negative effects on complexity or content and organization. This indicates that the improvement in accuracy was not at the cost of other aspects of writing. Such findings refute the predictions stipulated in the Trade-off Hypothesis (see Section 2.3) and, in a way, enrich our theoretical understanding.

This study has implications for L2 writing instruction, which may contribute to L2 learners' sustainable improvement of their writing performance. Firstly, it lends support to the value and usefulness of comprehensive WCF for L2 learners' writing performance in terms of writing accuracy as well as fluency. As such, it is reasonable that L2 teachers continue such an instructional practice in their normal writing pedagogy. Another implication is how to implement rewriting after the provision of direct WCF. As an important pedagogical practice, rewriting activity is worth carrying out after WCF provision, in which students have opportunities to practice and reinforce the knowledge they have learnt from WCF in revision. However, it would be of little use for students to copy their teachers' corrections after direct WCF in rewriting process. Alternatively, our study asked students to study the corrections but they did not have access to them in revision. While this practice is not common in real L2 writing classrooms, it enables learners to have a better understanding of WCF and their errors than just copying corrections passively. Our study probably provides teachers with a useful alternative, in which they can deliver their WCF directly (a ubiquitous practice in natural classroom settings) together with students' self-reflections while they revise their writing, which may maximize the efficacy of direct WCF and help foster self-sufficient revisers. Finally, as indicated by our results, WCF did not exert any effects on the content or the organization quality. Thus, to help students achieve a balanced development in L2 writing, teachers should not excessively place much emphasis on linguistic errors, at the expense of ignoring problems related to the content and organization of the written texts when providing feedback. This pedagogical suggestion based on our research results is in line with the recommendation that a balanced coverage on different aspects of writing should be stressed in feedback provision [82,83]. Focusing on both local and global issues in response to L2 learners' writing is believed to bring about the sustainable growth of their writing skills.

Author Contributions: X.C. and L.J.Z. conceived and designed the study. X.C. collected the data and drafted the manuscript, and all the authors revised the manuscript. L.J.Z. finalized it for submission as the corresponding author. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study involving human participants was reviewed and approved by the Human Participants Ethics Committee of The University of Auckland (020361, November 2017).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to ethical considerations.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Pretest:

Nowadays, our life is getting a lot simpler and more convenient because of various intelligent machines. However, some people think that our brains will get lazy in a world run by intelligent machines. Write a composition of at least 200 words on the following topic: *With intelligent machines to do the thinking, will our brains get lazy?*

Posttest:

An undergraduate of English at a university, in a recent letter to the university's president, complained that the mandatory math classes he had to take. He said that because a language major has little use for math, he would forget all of his math lessons soon after taking the required exams. Write a composition of at least 200 words on the following topic: *Should university students of English major learn math?*

Delayed posttest:

Recently, a survey reported that 67% of Chinese university students think that saving money is a good habit while the rest believed that spending tomorrow's money today is better. Write a composition of at least 200 words on the following topic: *Should university students save money or spend tomorrow's money?*

References

1. Manchon, R. *Learning-to-Write and Writing-to-Learn in an Additional Language*; Benjamins: Amsterdam, The Netherlands, 2011.
2. Zhang, L.J. Second language writing as and for second language learning. *J. Second. Lang. Writ.* **2013**, *22*, 446–447. [[CrossRef](#)]
3. Zhang, L.J. Reflections on the pedagogical imports of western practices for professionalizing ESL/EFL writing and writing-teacher education. *Aust. Rev. Appl. Linguist.* **2016**, *39*, 203–232. [[CrossRef](#)]
4. Cheng, X.; Zhang, L.J. Teacher Written Feedback on English as a Foreign Language Learners' Writing: Examining Native and Nonnative English-Speaking Teachers' Practices in Feedback Provision. *Front. Psychol.* **2021**, *12*, 629921. [[CrossRef](#)] [[PubMed](#)]
5. Lee, I. Utility of focused/comprehensive written corrective feedback research for authentic L2 writing classrooms. *J. Second. Lang. Writ.* **2020**, *49*, 100734. [[CrossRef](#)]
6. Li, H.H.; Zhang, L.J.; Parr, J.M. Small-Group Student Talk Before Individual Writing in Tertiary English Writing Classrooms in China: Nature and Insights. *Front. Psychol.* **2020**, *11*, 570565. [[CrossRef](#)]
7. Truscott, J. The Case Against Grammar Correction in L2 Writing Classes. *Lang. Learn.* **1996**, *46*, 327–369. [[CrossRef](#)]
8. Truscott, J. The effect of error correction on learners' ability to write accurately. *J. Second. Lang. Writ.* **2007**, *16*, 255–272. [[CrossRef](#)]
9. Truscott, J. The Effectiveness of Error Correction: Why do Met-Analytic Reviews Produce Such Different Answers. In *Epoch Making in English Teaching and Learning*; Leung, Y., Ed.; Crane: Taipei, Taiwan, 2016; pp. 126–141.
10. Bitchener, J.; Knoch, U. The value of a focused approach to written corrective feedback. *ELT J.* **2009**, *63*, 204–211. [[CrossRef](#)]
11. Ellis, R.; Sheen, Y.; Murakami, M.; Takashima, H. The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System* **2008**, *36*, 353–371. [[CrossRef](#)]
12. Ferris, D.R. The "Grammar Correction" Debate in L2 Writing: Where are we, and where do we go from here? (and what do we do in the meantime . . . ?). *J. Second. Lang. Writ.* **2004**, *13*, 49–62. [[CrossRef](#)]
13. Karim, K.; Nassaji, H. The revision and transfer effects of direct and indirect comprehensive corrective feedback on ESL students' writing. *Lang. Teach. Res.* **2020**, *24*, 519–539. [[CrossRef](#)]
14. Shintani, N.; Aubrey, S. The effectiveness of synchronous and asynchronous written corrective feedback on grammatical accuracy in a computer-mediated environment. *Mod. Lang. J.* **2016**, *100*, 296–319. [[CrossRef](#)]
15. Nassaji, H. Assessing the effectiveness of interactional feedback for L2 acquisition: Issues and challenges. *Lang. Teach.* **2020**, *51*, 3–28. [[CrossRef](#)]
16. Lee, I.; Mak, P.; Burns, A. Bringing innovation to conventional feedback approaches in EFL secondary writing classrooms: A Hong Kong case study. *Engl. Teach. Pract. Crit.* **2015**, *14*, 140–163. [[CrossRef](#)]
17. Rummel, S.; Bitchener, J. The effectiveness of written corrective feedback and the impact Lao learners' beliefs have on uptake. *Aust. Rev. Appl. Linguist.* **2015**, *38*, 66–84. [[CrossRef](#)]
18. Sheen, Y.; Wright, D.; Moldawa, A. Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System* **2009**, *37*, 556–569. [[CrossRef](#)]
19. Evans, N.W.; James Hartshorn, K.; Strong-Krause, D. The efficacy of dynamic written corrective feedback for university matriculated ESL learners. *System* **2011**, *39*, 229–239. [[CrossRef](#)]
20. Hartshorn, K.J.; Evans, N.W.; Merrill, P.F.; Sudweeks, R.R.; Strong-Krause, D.; Anderson, N.J. Effects of Dynamic Corrective Feedback on ESL Writing Accuracy. *TESOL Q.* **2010**, *44*, 84–109. [[CrossRef](#)]
21. Van Beuningen, C.; De Jong, N.H.; Kuiken, F. The effect of direct and indirect corrective feedback on L2 Learners' written accuracy. *ITL Int. J. Appl. Linguist.* **2008**, *156*, 279–296. [[CrossRef](#)]
22. Van Beuningen, C.G.; De Jong, N.; Kuiken, F. Evidence on the Effectiveness of Comprehensive Error Correction in Second Language Writing. *Lang. Learn.* **2011**, *62*, 1–41. [[CrossRef](#)]
23. Zhang, T. The effect of highly focused versus mid-focused written corrective feedback on EFL learners' explicit and implicit knowledge development. *System* **2021**, *99*, 102493. [[CrossRef](#)]
24. Bitchener, J.; Knoch, U. Written corrective feedback studies: Approximate replication of Bitchener & Knoch (2010) and Van Beuningen, De Jong & Kuiken (2012). *Lang. Teach.* **2015**, *48*, 405–414.
25. Lee, I. *Classroom Writing Assessment and Feedback in L2 School Contexts*; Springer: Singapore, 2017.
26. Zhang, L.J. Appraising the role of written corrective feedback in EFL writing. In *Reconceptualizing English Language Teaching and Learning in the 21st Century*; Leung, Y.-N., Katchen, J., Hwang, S.-Y., Chen, Y., Eds.; Crane: Taipei, Taiwan, 2018; pp. 134–146.
27. Bitchener, J.; Storch, N. *Written Corrective Feedback for L2 Development*; Multilingual Matters: Bristol, UK, 2016.
28. Li, S.; Vuono, A. Twenty-five years of research on oral and written corrective feedback in System. *System* **2019**, *84*, 93–109. [[CrossRef](#)]
29. Liu, Y.; Zhang, L.J.; May, S. Dominance of Anglo-American cultural representations in university English textbooks in China: A corpus linguistics analysis. *Lang. Cult. Curric.* **2021**, 1–19. [[CrossRef](#)]
30. Rao, Z.; Li, X. Native and Non-native Teachers' Perceptions of Error Gravity: The Effects of Cultural and Educational Factors. *Asia-Pac. Educ. Res* **2017**, *26*, 51–59. [[CrossRef](#)]

31. Teng, L.S.; Zhang, L.J. Empowering learners in the second/foreign language classroom: Can self-regulated learning strategies-based writing instruction make a difference? *J. Second. Lang. Writ.* **2020**, *48*, 100701. [[CrossRef](#)]
32. Ashwell, T. Patterns of Teacher Response to Student Writing in a Multiple-Draft Composition Classroom: Is Content Feedback Followed by Form Feedback the Best Method? *J. Second. Lang. Writ.* **2000**, *9*, 227–257. [[CrossRef](#)]
33. Ferris, D.R.; Roberts, B. Error feedback in L2 writing classes: How explicit does it need to be? *J. Second Lang. Writ.* **2001**, *10*, 161–184. [[CrossRef](#)]
34. Bitchener, J.; Knoch, U. The value of written corrective feedback for migrant and international students. *Lang. Teach. Res.* **2008**, *12*, 409–431. [[CrossRef](#)]
35. Bitchener, J.; Knoch, U. The Contribution of Written Corrective Feedback to Language Development: A Ten Month Investigation. *Appl. Linguist.* **2010**, *31*, 193–214. [[CrossRef](#)]
36. Sheen, Y. The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Q.* **2007**, *41*, 255–283. [[CrossRef](#)]
37. Ferris, D.R. Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Stud. Second Lang. Acquis.* **2010**, *32*, 181–201. [[CrossRef](#)]
38. Ferris, D.R.; Liu, H.; Sinha, A.; Senna, M. Written corrective feedback for individual L2 writers. *J. Second. Lang. Writ.* **2013**, *22*, 307–329. [[CrossRef](#)]
39. Zhang, T.; Chen, X.; Hu, J.; Ketwan, P. EFL Students' Preferences for Written Corrective Feedback: Do Error Types, Language Proficiency, and Foreign Language Enjoyment Matter? *Front. Psychol.* **2021**, *12*, 660564. [[CrossRef](#)]
40. Van Beuningen, C. Corrective Feedback in L2 Writing: Theoretical Perspectives, Empirical Insights, and Future Directions. *Int. J. Engl. Stud.* **2010**, *10*, 1–27. [[CrossRef](#)]
41. Bitchener, J. Evidence in support of written corrective feedback. *J. Second. Lang. Writ.* **2008**, *17*, 102–118. [[CrossRef](#)]
42. Shintani, N.; Ellis, R. The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *J. Second. Lang. Writ.* **2013**, *22*, 286–306. [[CrossRef](#)]
43. Liu, Q.; Brown, D. Methodological synthesis of research on the effectiveness of corrective feedback in L2 writing. *J. Second. Lang. Writ.* **2015**, *30*, 66–81. [[CrossRef](#)]
44. Storch, N. Critical feedback on written corrective feedback research. *Int. J. Engl. Stud.* **2010**, *10*, 29–46. [[CrossRef](#)]
45. Brown, D. The Written Corrective Feedback Debate: Next Steps for Classroom Teachers and Practitioners. *TESOL Q.* **2012**, *46*, 861–867. [[CrossRef](#)]
46. Hartshorn, J.; Evans, N.W. The effects of dynamic written corrective feedback: A 30-week study. *J. Response Writ.* **2015**, *1*, 6–34.
47. Bonilla López, M.; Van Steendam, E.; Speelman, D.; Buyse, K. The Differential Effects of Comprehensive Feedback Forms in the Second Language Writing Class. *Lang. Learn.* **2018**, *68*, 813–850. [[CrossRef](#)]
48. Sun, S. Written corrective feedback: Effects of Focused and Unfocused Grammar Correction on the Case Acquisition in L2 German. Ph.D. Thesis, University of Kansas, Lawrence, KS, USA, 2013.
49. Truscott, J.; Hsu, A.Y.-P. Error correction, revision, and learning. *J. Second. Lang. Writ.* **2008**, *17*, 292–305. [[CrossRef](#)]
50. Kang, E.; Han, Z. The Efficacy of Written Corrective Feedback in Improving L2 Written Accuracy: A Meta-Analysis. *Mod. Lang. J.* **2015**, *99*, 1–18. [[CrossRef](#)]
51. Skehan, P. *A Cognitive Approach to Language Learning*; Oxford University Press: Oxford, UK, 1998.
52. Skehan, P. Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Appl. Linguist.* **2009**, *30*, 510–532. [[CrossRef](#)]
53. Rahimi, M.; Zhang, L.J. Effects of task complexity and planning conditions on L2 argumentative writing production. *Discourse Process.* **2018**, *55*, 726–742. [[CrossRef](#)]
54. Kepner, C. An experiment in the relationship of types of written feedback to the development of second language writing skills. *Mod. Lang. J.* **1991**, *75*, 305–313. [[CrossRef](#)]
55. Semke, H. The effects of the red pen. *Foreign Lang. Annu.* **1984**, *17*, 195–202. [[CrossRef](#)]
56. Chandler, J. The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *J. Second Lang. Writ.* **2003**, *12*, 267–296. [[CrossRef](#)]
57. Rahimi, M. The role of teacher's corrective feedback in improving Iranian EFL learners' writing accuracy over time: Is learner's mother tongue relevant? *Read. Writ.* **2008**, *22*, 219–243. [[CrossRef](#)]
58. Rahimi, M. A comparative study of the impact of focused vs. comprehensive corrective feedback and revision on ESL learners' writing accuracy and quality. *Lang. Teach. Res* **2019**, 1–24. [[CrossRef](#)]
59. Mao, Z.; Lee, I. Feedback scope in written corrective feedback: Analysis of empirical research in L2 contexts. *Assess. Writ.* **2020**, *45*, 100469. [[CrossRef](#)]
60. Guo, Q.; Barrot, J.S. Effects of Metalinguistic Explanation and Direct Correction on EFL Learners' Linguistic Accuracy. *Read. Writ. Q.* **2019**, *35*, 261–276. [[CrossRef](#)]
61. Dörnyei, Z. *Research Methods in Applied Linguistics*; Oxford University Press: Oxford, UK, 2007.
62. Bitchener, J.; Ferris, D.R. *Written Corrective Feedback in Second Language Acquisition and Writing*; Routledge: London, UK, 2012.
63. Ferris, D.R. *Response to Student Writing: Research Implications for Second Language Students*; Erlbaum: Mahwah, NJ, USA, 2002.
64. Ekanayaka, W.I.; Ellis, R. Does asking learners to revise add to the effect of written corrective feedback on L2 acquisition? *System* **2020**, *94*, 102341. [[CrossRef](#)]

65. Lu, X. Automatic analysis of syntactic complexity in second language writing. *Int. J. Corpus Linguist.* **2010**, *15*, 474–496. [[CrossRef](#)]
66. Lu, X. A Corpus-Based Evaluation of Syntactic Complexity Measures as Indices of College-Level ESL Writers' Language Development. *TESOL Q.* **2011**, *45*, 36–62. [[CrossRef](#)]
67. Rahimi, M.; Zhang, L.J. Writing task complexity, students' motivational beliefs, anxiety and their writing production in English as a second language. *Read. Writ.* **2019**, *32*, 761–786. [[CrossRef](#)]
68. Geng, H. The effects of written feedback and revision-focus direction on student revision and writing improvement. Doctoral Dissertation, The University of Auckland, Auckland, New Zealand, 2017. Unpublished work.
69. Jacobs, H.L.; Zinkgraf, S.A.; Wormuth, D.R.; Hartfiel, V.F.; Hughey, J.B. *Testing ESL Composition: A Practical Approach*; Newbury House: Rowley, MA, USA, 1981.
70. Fleiss, J.L.; Cohen, J. The Equivalence of Weighted Kappa and the Intraclass Correlation Coefficient as Measures of Reliability. *Educ. Psychol. Meas.* **1973**, *33*, 613–619. [[CrossRef](#)]
71. Field, A. *Discovering Statistics Using SPSS*; Sage: London, UK, 2009.
72. Cohen, J. Quantitative methods in psychology: A power primer. *Psychol. Bull.* **1992**, *112*, 155–159. [[CrossRef](#)] [[PubMed](#)]
73. Frear, D.; Chiu, Y.-H. The effect of focused and unfocused indirect written corrective feedback on EFL learners' accuracy in new pieces of writing. *System* **2015**, *53*, 24–34. [[CrossRef](#)]
74. Robinson, P. Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Appl. Linguist.* **2001**, *22*, 27–57. [[CrossRef](#)]
75. Robinson, P. *Second Language Task Complexity*; John Benjamins Publishing Company: Amsterdam, The Netherlands, 2011.
76. Khezrlou, S. The role of task repetition with direct written corrective feedback in L2 writing complexity, accuracy and fluency. *J. Second. Lang. Stud.* **2020**, *3*, 31–54. [[CrossRef](#)]
77. Storch, N.; Tapper, J. The impact of an EAP course on postgraduate writing. *J. Engl. Acad. Purp.* **2009**, *8*, 207–223. [[CrossRef](#)]
78. Barkaoui, K. What and when second-language learners revise when responding to timed writing tasks on the computer: The roles of task type, second language proficiency, and keyboarding skills. *Mod. Lang. J.* **2016**, *100*, 320–340. [[CrossRef](#)]
79. Chen, J.; Zhang, L.J. Assessing student-writers' self-efficacy beliefs about text revision in EFL writing. *Assess. Writ.* **2019**, *40*, 27–41. [[CrossRef](#)]
80. Lee, I. ESL learners' performance in error correction in writing: Some implications for teaching. *System* **1997**, *25*, 465–477. [[CrossRef](#)]
81. Nassaji, H.; Swain, M. A Vygotskian perspective on corrective feedback in L2: The effect of random versus negotiated help on the learning of English articles. *Lang. Aware.* **2000**, *9*, 34–51. [[CrossRef](#)]
82. Ferris, D.R. Responding to student writing: Teachers' philosophies and practices. *Assess. Writ.* **2014**, *19*, 6–23. [[CrossRef](#)]
83. Reichelt, M. Contextual factors impacting feedback practices for non-English L2 writing. *Foreign Lang. Ann.* **2019**, *52*, 744–752. [[CrossRef](#)]