English Language Teaching; Vol. 13, No. 7; 2020
ISSN 1916-4742 E-ISSN 1916-4750
Published by Canadian Center of Science and Education

Effects of Written Corrective Feedback:

A Synthesis of 10 Quasi-Experimental Studies

Albatool Abalkheel[1] & Tara Brandenburg[2]

[1] Department of English Language and Translation, College of Humanities and Arabic Studies, Buraydah, Qassim University, Saudi Arabia

[2] College of Western Idaho, USA

Correspondence: Albatool Abalkheel, Department of English Language and Translation, College of Humanities and Arabic Studies, Buraydah, Qassim University, Saudi Arabia

Received: May 10, 2020 Accepted: June 23, 2020 Online Published: June 25, 2020

doi: 10.5539/elt.v13n7p97 URL: https://doi.org/10.5539/elt.v13n7p97

Abstract

Many language teachers spend countless hours correcting student writing in hopes of improvement in accuracy, but as of yet, there has been little consensus regarding the efficacy of written corrective feedback (CF) or the type of CF that is most efficient. Although many studies have been conducted on the topic, conflicting results have arisen. In this meta-analysis, ten quasi-experimental studies of written corrective feedback are examined to analysis the overall effect of CF and compare the variations of CF. It is shown that written corrective feedback in general is inconclusive as a predictor of student improvement in writing over time and the efficacy of the feedback depends on its focus. It is also shown that focused written feedback has any overall positive effect on student's writing, whereas comprehensive written feedback has the potential to have a harmful effect on student's writing over time.

Keywords: Direct written corrective feedback, indirect written corrective feedback, Quasi-Experimental Studies

1. Introduction

Over the years, many scholars have either taken strong opposition or support of corrective feedback (CF). On one, Truscott presents evidence that CF does not have any significant impact on long-term learning (Truscott & Hsu, 2008) and claims that it may in fact have a negative effect (Truscott, 2007), particularly if students refrain from experimenting with new vocabulary or forms in writing in anticipation of CF. Karim and Nassaji (2018) also find that CF just has short-term transfer effects on grammatical accuracy. They claim that it did not have any significant delayed transfer effects. Others claim to have empirically demonstrated the positive effects of CF on long term knowledge retention (Bitchner, 2008; Bitchner & Knoch, 2010; Sheen, 2009). However, among those who are in favor of CF, the questions still remain whether to use CF focused to particular selected error types or comprehensive CF targeting all errors. There is also some contention as to whether direct (provision of the correct form) or indirect (indication of an error only) CF is most effective. Although numerous studies have been conducted on the subject, recent research lacks a complete synthesis of the effects of written corrective feedback. Hence, the following research questions will be investigated:

1. Does written corrective feedback (CF) increase accuracy in subsequent writing?

2. What type of CF is the most effective?

2. Method

A review of quasi-experimental studies regarding written corrective feedback was first conducted. Only articles published in major journals of linguistics (e.g. *TESOL Quarterly*, *Language Learning*, *Journal of Second Language Writing, System*, *Applied Linguistics*, *Language Teaching Research*, *The Modern Language Journal*) were considered to be potential sources. The parameters of studies selected included, (1) that the study must employ a pre- and a post-test, (2) studies must have been published after 1990, and (3) studies must be limited to the effects of written corrective feedback.

The methodologies utilized in the selected studies were similar. All studies were conducted within in-tact second or foreign language writing classes. The pre-tests consisted of students responding to a writing prompt. The

writing was scored as a pre-test. The treatment of corrective feedback (or lack thereof) was applied to the initial writing sample and any subsequent writing samples within the treatment period. All groups (treatment and control) were instructed in writing and grammar throughout the treatment period and all groups within each study wrote the same number of essays, but received different feedback or lack thereof. The post-test for each group consisted of another writing prompt that was scored and recorded by the researchers based on the same criteria as the pre-test.

2.1 Focus of Feedback

The primary distinction in the focus of feedback provided was comprehensive or focused feedback. Comprehensive feedback refers to feedback in which the examiner corrects every error. Focused feedback is when the examiner determines certain types of errors before examining the writing and then only responds to errors pertaining to the selected error categories. One study compared comprehensive and focused feedback (Sheen et al., 2009), while four studies measured the effects of comprehensive feedback (Lalande, 1982; Polio et al., 1998; Truscott & Hsu, 2008; Hartshorn et al., 2010; Van Beuningen et al. 2011), and the remaining four utilized focused feedback that only targeted particular grammatical features (Bitchner & Knoch, 2008; Bitchner, 2008; Bitchner & Knoch, 2009; Bitchner & Knoch, 2010). Of the focused feedback studies, all feedback pertained to the use of referential 'a' and 'the'.

2.2 Type of Feedback

The selected studies distinguish between direct written corrective feedback (DCF) or indirect written corrective feedback (ICF). DCF is defined as feedback that "provides some form of explicit correction of linguistic form or structure above or near the linguistic error. It may consist of the crossing out of an unnecessary word/phrase/morpheme, the insertion of a missing word/phrase/morpheme, and the provision of the correct form or structure" (Bitchner & Knoch, 2010, p. 209). ICF on the other hand, indicates the presence of an error, but does not provide the correct form. Some common ways of providing ICF are through circling or underlining errors, writing the number of errors in each line in the margin, or using meta-linguistic codes to indicate the type of error. ICF has been argued to more deeply engage students cognitively than feedback in which the correct form is supplied (Bitchener, 2008).

One study tested the efficacy of an unspecified class of CF versus the lack of CF (Polio et al., 1998), whereas the others studied all specified the kind of written corrective feedback. ICF versus no CF was tested in three of the studies (Lalande, 1982; Truscott & Hsu, 2008; Hartshorn et al., 2010). DCF was tested in comparison to no CF in four of the studies (Bitchner & Knoch, 2008; Bitchner, 2008; Bitchner & Knoch, 2009, Sheen et al., 2009) and DCF was compared to ICF in two studies (Bitchner & Knoch, 2010; Van Beuningen et al., 2011). For control groups in which no CF was provided, participants were provided with multiple practice writing assignments and writing and grammar instruction.

Table 1. Overview of Ten Selected Studies

	Participants	Setting	Target forms	Feedback type	Length
Lalande (1982)	60 intermediate L2 German students	university in the US	comprehensive	1. ICF + error awareness sheet 2.C- CF *note, described as "traditional," but not exemplified	1 quarter
Polio et al. (1998)	64 advanced ESL students	university in the US	all grammatical errors	1. CF + grammar exercises 2. C- no CF	semester
Truscott & Hsu (2008)	47 EFL graduate students	Taiwan	all grammatical errors	1. ICF 2. C- no CF	2 weeks
Bitchner & Knoch (2008)	144 intermediate	university in New Zealand	referential 'a' and 'the'	1. DCF + written & oral metalinguistic exp.	2 months

	ESL students			2. DCF + written metalinguistic exp. 3. DCF 4. C – no CF	
Bitchner (2008)	75 intermediate ESL students	university in New Zealand	referential ‘a’ and ‘the’	1. DCF + written & oral metalinguistic exp. 2. DCF + written metalinguistic exp. 3. DCF 4. C – no CF	2 months
Bitchner & Knoch (2009)	52 intermediate ESL students	university in New Zealand	referential ‘a’ and ‘the’	1. DCF + written & oral metalinguistic exp. 2. DCF + written metalinguistic exp. 3. DCF 4. C – no CF	10 weeks
Sheen et al. (2009)	80 intermediate ESL students	college in the USA	referential ‘a’ and ‘the’ vs. comprehensive (unfocused)	1. Focused DCF 2. Unfocused DCF 3. C- No CF + writing practice 4. C – no writing; grammar practice only	9 weeks
Hartshorn et al. (2010)	47 advanced ESL students	university in the USA	comprehensive	1.ICF + tally sheets and error lists 2. C- ICF only	15 weeks
Bitchner & Knoch (2010)	63 advanced ESL students	university in the USA	referential ‘a’ and ‘the’	1. DCF 2. ICF 3. DCF + oral metalinguistic exp. 4. C- no CF	10 weeks
Van Beuningen et al. (2011)	268 L2 Dutch students	Multilingual students of four Dutch secondary schools	comprehensive	1. DCF 2. ICF 3. C- Self Correction, no CF 4. C- additional practice, no CF	6 weeks

Note. ** CF- Corrective Feedback; DCF- Direct Corrective Feedback; ICF- Indirect Corrective Feedback; C- Control

2.3 Outcome Measures

To measure the effect of the treatment, pre- and post-tests were administered in all studies. In all studies, the measure used for the pre-test was equivalent to the measure used in the post-test. The pre-tests of the studies generally demonstrated the equivalency of the test and the control groups. Hence, only the post-tests scores were utilized for in the meta-analysis. In the case that both a post-test and a delayed post-test were used as a measure after the intervention, only the delayed post-test results were included in the effect size calculation because the goal of this study is to determine the effects of CF on acquisition rather than just short-term knowledge.

While all of the studies utilized writing samples as the outcome measure, there was some variation in how the outcome measures were scored (See Table 2). Two counted the number of grammatical errors in writing (Lalande, 1982; Truscott & Hsu, 2008), one utilized error free t-units divided by total t-units in the writing sample (Polio et al., 1998), five utilized correct use of targeted features in writing (Bitchner & Knoch, 2008; Bitchner, 2008; Bitchner & Knoch, 2009; Sheen et al., 2009; Bitchner & Knoch, 2010), one utilized final essay ratings based on a rubric (Hartshorn et al., 2010) and one utilized overall accuracy of writing (Van Beuningen et al., 2011).

Not only were different scoring systems utilized, a variety of non-comparable statistics were used to calculate the outcome of the procedure. Hence, it is necessary to convert the scores into a comparable format. This was done by calculating Cohen's *d* to examine the effect size. This can be calculated with the standard deviation and means of the treatment and control groups. In this meta-analysis, Cohen's *d* was calculated using the *Effect Size Calculator* (Becker, 1998).

3. Result and Discussion

For the interpretation of this meta-analysis, the conventional benchmarks for interpreting effect size with Cohen's *d* will be used, which are small at r=.10 (explains 1% of variance), medium at r = .30 (explains 9% of variance), and large at r=.50 (explains 25% of variance) (Field & Gillett, 2010, p. 669). It will also be necessary to note that "the difference is positive if it is in the direction of *improvement* or in the predicted direction and negative if in the direction of *deterioration* or opposite to the predicted direction" (Becker, 2000).

3.1 Effects of Written Corrective Feedback

Results of effect size calculation (See Table 2) demonstrated large effect sizes in five of the studies (Bitchner & Knoch, 2008; Bitchner, 2008; Bitchner & Knoch, 2009; Harshorn et al., 2010; Bitchner & Knoch, 2010). Two of the studies did not demonstrate significant effect sizes (Polio et al., 1998; Truscott & Hsu, 2008). One study generated mixed results with a large positive effect size for the Focused DCF group, whereas a large negative effect size was demonstrated for the Unfocused DCF group (Sheen et al., 2009). Only one study demonstrated an overall negative effect size (Van Bueningen et al., 2011) with a small negative effect with the high-DCF group and the low- ICF group, a medium negative effect for the low-DCF group and a large negative effect size for the high-ICF. Although the results of Lalande (1982) appeared significant, it was excluded from the analysis of the effect of CF in general because the control group received CF.

With five studies clearly demonstrating a large positive effect size, two neutral, one mixed and one negative, it appears that in about 55% the cases, CF has a positive effect on students' long-term learning. On the other hand, there is about a 22% chance that CF will have no significant effect at all on students, and at worse, it appears that there is about a 16% chance that CF will have a negative effect on long term learning, causing accuracy to actually decrease over time. However, as the mixed results of the Sheen et al. (2009) study reveal, the type of CF provided can be the deciding variable that determines whether or not CF will be beneficial or harmful to students.

Table 2. Summary of effect size

Study	Groups	*n*	Mean	S.D.	Effect sizes (Cohen's d)	Measure
Lalande (1982)	ICF + error awar.	30	25.23	14.19	-0.512	total of non-lexical errors in writing
	C- CF only	30	32.87	15.61		
Polio et al. (1998)	CF+grammar ex.	31	.235	1.44	-0.0136	error free t-units /total t-units in writing
	C- no CF	34	.249	0.127		
Truscott & Hsu	ICF	21	.1130	.0472	0.0692	number of grammatical errors

(2008)	C- no CF	26	.1095	.0537		in writing
Bitchner & Knoch (2008)	DCF+written&oral	34	80.91	12.80	**1.1548	Correct use of targeted features in writing
	DCF+written	36	76.28	16.45	**0.8053	
	DCF	35	79.20	12.92	**1.0499	
	C-no CF	39	61.56	19.94		
Bitchner (2008)	DCF+written&oral	17	82.65	15.47	**1.0856	correct use of targeted features in writing
	DCF+written	18	76.78	20.67	**0.6503	
	DCF	20	80.05	12.10	**1.0177	
	C-no CF	20	63.90	18.90		
Bitchner & Knoch (2009)	DCF+written&oral	13	88.38	9.53	**2.2207	correct use of targeted features in writing
	DCF+written	13	88.77	8.55	**2.3090	
	DCF	13	81.46	13.90	**1.4954	
	C-no CF	13	58.92	16.16		
Sheen et al. (2009)	Focused DCF	22	78	13.9	*0.3211	measured by accuracy of a/the in writing
	Unfocused DCF	23	67	17.7	-0.4473	
	C- No CF + writing practice	16	73.8	12.2		
Hartshorn et al. (2010)	ICF + tally/error sheet	28	24.16	19.46	**0.6448	final essays ratings
	C-ICF only	19	13.78	11.81		
Bitchner & Knoch (2010)	DCF	12	96.25	4.00	**1.1662	measured by accuracy of a/the in writing
	ICF	27	92.76	6.80	**0.7071	
	DCF + oral	12	95.33	5.09	**1.0303	
	C- no CF	12	85.92	11.87		
Van Beuningen et al. (2011)	High level					overall accuracy of writing
	DCF	32	1.32	0.89	-0.2248	
	ICF	29	0.89	0.40	-1.0700	
	C- SLF	37	1.50	0.70		
	Low level					
	DCF	35	1.42	0.78		
	ICF	33	1.54	0.90	-0.4071	
	C- SLF	34	1.75	0.84	-0.2412	

Note: in groups with multiple controls, only the first control group was used in calculating the effect size.

3.2 Effects of Type of Written Corrective Feedback

The results of the more efficient type of CF proved to be inconclusive for this group of studies. For DCF, five studies displayed a large positive effect size and two were negative. In other words, there is a 29% chance that DCF could actually cause students work to worsen over time. For ICF, the results were as conclusive as a coin flip. Two studies displayed large positive effect size, one was insignificant, and two were negative (See Table 3). Hence, it is not possible to determine whether DCF of ICF is more efficient without examining other variables.

Table 3. Effect Size by Variation in CF Type and Focus

Type of CF by Study	Effect Size	*n*	Focus of CF
DCF			
Bitchner & Knoch (2008)	+ large	144	focused
Bitchner (2008)	+ large	75	focused
Bitchner & Knoch (2009)	+ large	52	focused
Sheen et al. (2009)	+ medium	22	focused
	- large	23	comprehensive
Bitchner & Knoch (2010)	+ large	24	focused
Van Beuningen et al. (2011) -high level	- small	32	comprehensive
	- large	35	comprehensive
-low level			
ICF			
Lalande (1982)	- large	60	comprehensive
Truscott & Hsu (2008)	insignificant	47	comprehensive
Hartshorn et al. (2010)	+ large	47	comprehensive
Bitchner & Knoch (2010)	+ large	27	focused
Van Beuningen et al. (2011) -high level	- large	29	comprehensive
	- small	29	comprehensive
-low level			

Note: Polio et al. was excluded from this analysis because CF type was unspecified

What proved to be more telling of the efficacy of feedback was the focus, rather than the type of feedback. In all cases, focused feedback had a positive effect size in both DCF and ICF. However, it should be noted that more research needs to be done to determine the efficacy of focused ICF because this was only supported by one study (Bitchner & Knoch, 2010). On the other hand, the results for comprehensive feedback demonstrated negative effect sizes for all DCF studies and mixed, with half negative (Lalande, 1982; VanBueningen et al., 2011) and half positive (Truscott & Hsu, 2008; Hartshorn et al., 2010) for comprehensive ICF studies.

4. Conclusion

Written corrective feedback, either direct or indirect, in general is inconclusive as a predictor of student improvement in writing over time. Rather than the presence or absence of feedback, the efficacy of the feedback depends on its focus. It was shown that focused written feedback has an overall positive effect on student's writing, whereas comprehensive written feedback has the potential to have a harmful effect on student's writing over time.

Sheen et al. suggest that the reason why focused instruction is more effective than comprehensive is that "when the correction addresses a range of grammatical errors, learners are unable to process the feedback effectively, and even if they attend to the corrections, they are unable to work out why they have been corrected" (2009, pp. 565-566). Furthermore, they argue that the range of errors attended to in comprehensive feedback may overburden students and that they are often unsystematic and arbitrarily selected, whereas focused CF helps leaners to notice errors, systematically engage in hypothesis testing, and monitor their own writing through the use of existing grammatical knowledge.

The pedagogical implications of this meta-analysis point to the correction methodologies described by Bitchner (2008), Bitchner & Knoch (2008, 2009, 2010) and Sheen (2009). Rather than addressing all errors in a composition, the instructor should select a limited number of error types and only address those in written corrective feedback.

While the error type was limited to two categories in these studies, Bitchner and Knoch suggest that the number could potentially be increased, particularly with advanced learners (2010). Bitchner and Knoch recommend that "the provision of clear, simple meta-linguistic explanation, namely, explanation of rule(s) with example(s), is the best type of written CF for long-term accuracy" (2010, p. 216). However, due to the limited number of studies included in this meta-analysis and the potential confounding variable of including four studies by the same principle researcher, more research should be conducted in order to determine more conclusive results.

References

Becker, L.A. (1998). *Effect Size Calculators.* Retrieved from http://www.uccs.edu/~lbecker/.

Becker, L.A. (2000). *Standardized difference between two groups.* Retrieved from *http://www.uccs.edu/lbecker/effect-size.html#V*

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

Bitchner, J. & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research, 12*(3), 409-431. https://doi.org/10.1177/1362168808089924

Bitchner, J. & Knoch, U. (2009). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics, 31*(2), 193-214. https://doi.org/10.1093/applin/amp016

Bitchner, J. & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing, 19,* 207-217. https://doi.org/10.1016/j.jslw.2010.10.002

Coe, R. (2002). It's the effect size, stupid. What effect size is and why it is important. Retrieved from: http://www.leeds.ac.uk/educol/documents/00002182.htm .

Field, A. P. & Gillett, R. (2010). Expert tutorial: How to do a meta-analysis. *The British Psychological Society, 63,* 665-694. https://doi.org/10.1348/000711010X502733

Hartshorn, K.J., Evans, N.W., Merril, P.F., Sudweeks, R.R., Strong-Krause, D., & Anderson, N.J. (2010). Effects of dynamic corrective feedback on ESL writing accuracy. *TESOL Quarterly, 44*(1), 84-109. https://doi.org/10.5054/tq.2010.213781

Karim, K. and Nassaji, H. (2018) The revision and transfer effects of direct and indirect comprehensive corrective feedback on English-as-a-second-language (ESL) students' writing. *Language Teaching Research*. https://doi.org/10.1177/1362168818802469

Polio, C., Fleck, C. & Leder, N. (1998). "If I only had more time:" ESL learners' changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing, 7*(1), 43-68. https://doi.org/10.1016/S1060-3743(98)90005-4

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,* 556-569. https://doi.org/10.1016/j.system.2009.09.002

Truscott, J. (2007). The effect of error correction on learns' ability to write accurately. *Journal of Second Language Writing, 16,* 255-272. https://doi.org/10.1016/j.jslw.2007.06.003

Truscott, J. & Hsu, A. Y. (2008). Error correction, revision, and learning. *Journal of Second Language Writing, 17,* 292-305. https://doi.org/10.1016/j.jslw.2008.05.003

Van Beuningen, C.G., De Jong, N.H., & Kuiken, F. (2011). Evidence of the effectiveness of comprehensive error correction in second language writing. *Language Learning: A Journal of Research in Language Studies, 62*(1), 1-41. https://doi.org/10.1111/j.1467-9922.2011.00674.x

Copyrights