Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 17. Number 3, September 2026 Pp.42-57
DOI: https://dx.doi.org/10.24093/awej/vol17no3.3

Unfocused Corrective Feedback and L2 Writing CAF: The Moderating Role of Individual Differences

Laurence Craven
American University of Sharjah, Department of English, Sharjah,
United Arab Emirates
Corresponding Author: lcraven@aus.edu

Nasser Jabbari
University of Essex, School of Language, Literature, and Media, UK
Email: n.jabbari@essex.ac.uk

Received:06/11/2026 Accepted: 08/20/2026 Published: 09/10/2026

Abstract
Research on written corrective feedback in second language learning has advanced, addressing uncertainties about its effectiveness. The debate now centres on the most effective way to give feedback. However, controversy persists, and research on unfocused feedback remains scarce. This quasi-experimental study examines the effects of unfocused written corrective feedback—direct, indirect, and metalinguistic—on the complexity, accuracy, and fluency of 139 English-as-a-second-language learners' writing at an English-medium instruction university in the United Arab Emirates. It also investigates the moderating effects of individual-difference variables—language aptitude, attitudes towards written corrective feedback, and second-language proficiency—on feedback efficacy. Three experimental groups and one control group wrote essays (pre-test), revised their drafts (post-test), and composed essays on new topics (delayed post-test) over fourteen weeks. MANCOVA results revealed that participants who received direct and metalinguistic feedback showed reduced fluency compared with those in the indirect feedback and control groups. Lexical diversity was significantly lower in the indirect feedback group than in the direct and metalinguistic groups. Unfocused feedback did not yield significant gains on new writing tasks, and none of the individual-difference variables significantly moderated the effects of feedback on complexity, accuracy, and fluency measures. These findings highlight differential short-term effects of unfocused written corrective feedback on specific writing dimensions, while suggesting limited transfer to new writing tasks and limited moderating influence of learner-level variables. The results provide useful information for practitioners, who could adopt a more blended approach between focused and unfocused feedback and increase the number of treatment sessions.

***Keywords*:** Accuracy, aptitude, complexity, fluency, individual differences, second language writing, unfocused written corrective feedback, United Arab Emirates

How to cite: Craven, L. & Jabbari, N. (2026). Unfocused Corrective Feedback and L2 Writing CAF: The Moderating Role of Individual Differences. *Arab World English Journal*, *17*(3): 42-57. https://dx.doi.org/10.24093/awej/vol17no3.3

Introduction

Over the past few decades, approaches to teaching English have been constantly evolving. Teaching academic writing to English as a second language (ESL) learners has also progressed. Despite these changes, one element remains constant: the incorporation of Written Corrective Feedback (WCF) in its three primary forms—direct, indirect, and metalinguistic (Ellis, 2008). Several researchers have found positive effects of WCF on students' writing and language development (e.g., Karim & Nassaji, 2020; Nicolás-Conesa et al., 2019; Bonilla López et al., 2018; Van Beuningen et al., 2012). Truscott's (1996) critique, however, sparked an ongoing debate about the effectiveness of WCF for ESL learners' writing. Truscott offered several theoretical and practical reasons for abandoning WCF. For example, he argued that simple information transfer through WCF could be ineffective, since language learning is gradual and does not merely involve a linear transfer of knowledge from teacher to student. He further contended that WCF may not align well with learners' developmental stages and that any learning resulting from WCF is likely "pseudo-learning" (Truscott, 1996, p. 345). Truscott's theoretical arguments have been countered by several researchers (e.g., Schmidt, 1990; Ellis, 2010; Bitchener & Ferris, 2012), who argue that, despite Truscott's (1996) assertion that simple information transfer through WCF may not be effective for knowledge acquisition, skill acquisition theories suggest that WCF can be beneficial, as the declarative knowledge it provides can transform into procedural knowledge through practice and gradual automatization (Bitchener & Ferris, 2012). Recent syntheses further confirm that WCF continues to play a significant role in classroom writing instruction, while highlighting persistent variation in learner engagement and outcomes (Mao et al., 2024).

Most WCF studies have examined group averages without considering individual student profiles. Calls have been made for research exploring how Individual Difference (ID) variables, such as Second Language (L2) proficiency, attitudes towards WCF, and language aptitude, moderate WCF effectiveness (Li, 2015; Kang & Han, 2015). Additionally, several studies (e.g., Bitchener & Storch, 2016; Ellis et al., 2008; Van Beuningen, 2010; Van Beuningen et al., 2012) emphasise the need to make research more applicable to classroom settings. Key recommendations include: (a) evaluating student improvement by comparing initial writing performance with subsequent writing on new topics, rather than simply revising original tasks; and (b) examining the effects of unfocused WCF, the most common form of feedback provided by language teachers.

Studies investigating all three types of unfocused WCF using complexity, accuracy, and fluency (CAF) measures are scarce. Karim and Nassaji (2020) focused exclusively on accuracy; Bonilla López et al. (2018) concentrated solely on accuracy measures; and Van Beuningen et al. (2012) assessed only accuracy and complexity. A large gap also remains regarding the moderating effects of ID variables on the efficacy of unfocused WCF when evaluated through CAF (Suzuki et al., 2019; Li, 2015). Despite a growing body of recent research, scholars continue to call for more classroom-based investigations examining how WCF types interact with learner-internal variables over time (Mao et al., 2024).

Current literature indicates that L2 proficiency moderates the effectiveness of WCF, with larger effect sizes at higher proficiency levels and negative effects for beginner learners (Kang & Han, 2015; Bitchener, 2008). Learners' attitudes towards feedback can influence whether and how they engage with corrections (Oxford & Shearin, 1994; Ellis, 2008). Researchers have also found that students with higher language aptitude benefit more from WCF (Granena & Long, 2013; Li, 2015); however, findings on which WCF type is most effective at specific levels of aptitude remain inconclusive.

Using a quasi-experimental design, this study aims to examine the effects of direct, indirect, and metalinguistic unfocused WCF on English-as-a-second-language (ESL) learners' written performance, using CAF indices, and to explore the moderating effects of language aptitude, attitudes, and L2 proficiency. This study's significance lies in its contributions to the WCF literature in several ways. Theoretically, it extends the debate between Truscott (1996) and skill acquisition theorists (e.g., Bitchener & Ferris, 2012) by examining whether unfocused WCF facilitates the transformation of declarative knowledge into procedural knowledge across multiple CAF dimensions, rather than focusing on accuracy alone. Methodologically, it responds to calls for more ecologically valid designs (Bitchener & Storch, 2016; Van Beuningen, 2010) by comparing students' performance on new writing tasks rather than on revisions of the same text, while accounting for ID variables (i.e., language aptitude, attitudes, and L2 proficiency) as moderators. In addition, the findings are expected to inform WCF practices in Gulf English-Medium Instruction (EMI) and similar contexts, where teacher-centred instruction remains prevalent (Chaudhary, 2022) and empirical guidance on which WCF type best supports writing development across proficiency and aptitude levels remains limited.

Aligned with the research questions below, this study aims to (a) examine the effects of direct, indirect, and metalinguistic unfocused WCF on CAF indices in ESL learners' revised writing (post-test), compared with no feedback, while accounting for ID variables as covariates; (b) examine the effects of the three WCF types on CAF indices in ESL learners' new writing (delayed post-test), compared with no feedback, while accounting for ID variables; and (c) determine the extent to which ID variables—language aptitude, attitudes, and L2 proficiency—affect writing performance outcomes after controlling for feedback type. Therefore, it seeks to address the following research questions:

RQ1: How do the three types of WCF affect CAF indices in students' revised writing (post-test) compared with no feedback, while accounting for ID variables as covariates?

RQ2: How do the three types of WCF influence CAF indices in students' new writing (delayed post-test) compared with no feedback, while accounting for ID variables?

RQ3: To what extent do ID variables affect writing performance outcomes after controlling for feedback type?

Literature Review

Focused and Unfocused WCF

The debate over which type of WCF is more effective continues. Ellis et al. (2008) and Sheen et al. (2009) found that learners notice and acquire a language form more effectively when WCF targets a single feature. Kang and Han's (2015) meta-analysis confirmed a much larger effect for focused than unfocused feedback. Frear and Chiu (2015), however, found focused and unfocused WCF equally effective at both post-test and delayed post-test, and Van Beuningen et al. (2012) noted that unfocused WCF may overwhelm students with too many corrections. Karim and Nassaji (2020) argued that focused WCF studies are not ecologically valid, since real classroom teachers typically correct a wide range of errors at once. Studies on unfocused WCF in more naturalistic conditions (Nicolás-Conesa et al., 2019; Karim & Nassaji, 2020; Bonilla López et al., 2018; Van Beuningen et al., 2012) show a consistent pattern: accuracy improves on revision tasks in the short term, but these gains do not transfer to new writing tasks. Kang and Han (2015) attribute the mixed results in this literature to the number of uncontrolled variables that accompany feedback type, including learner proficiency, task type, and treatment length. Razmi and Ghane (2024) found that

unfocused WCF affects not only writing performance but also learner self-efficacy and writing anxiety. Hartshorn and Pack (2026) found that AI-generated dynamic WCF improved accuracy but had inconsistent effects on complexity and fluency, the same pattern reported in earlier unfocused WCF research. Why unfocused WCF's short-term gains fail to transfer to new tasks remains unclear.

Direct, Indirect, and Metalinguistic WCF

The debate over which feedback form yields the best results persists. Some researchers favour direct WCF, arguing that it provides the correct form immediately and reduces cognitive demand (Van Beuningen et al., 2008; Bitchener & Knoch, 2010). Others, however, contend that indirect WCF is more effective because it requires learners to engage more deeply with their errors. Lalande (1982), for instance, argues that indirect WCF "[…] can invoke problem-solving/active-correction strategies" (p. 147), leading to better long-term writing development. In the same vein, Ferris (2006) asserts that indirect feedback prompts learners to seek out and construct the correct form themselves rather than merely noticing a given correction, leading to more robust and durable learning — a point often invoked to support the superiority of indirect WCF. More recently, Westmacott (2017) has similarly found that indirect WCF promotes deeper engagement with errors, lending further support to this line of argument. Metalinguistic feedback, which provides explicit information about the nature of errors, has received comparatively little research attention as an independent treatment condition despite its common use in classrooms. Almohawes (2025) found that learners preferred direct WCF on quantitative measures, yet qualitative interviews revealed a preference for metalinguistic feedback as a more engaging and memorable experience. Learner preferences in Arab Gulf EFL contexts tend towards more explicit feedback forms, particularly when writing confidence is limited (Mahmood, 2021; Chaudhary, 2022). No study has tested metalinguistic WCF as an independent condition in an Arab Gulf sample.

Individual Differences

Very few studies have explored and validated the role of L2 proficiency as a moderator of WCF effects (Ellis, 2009). Low-proficiency learners may prefer and benefit from focused, direct WCF, as it reduces cognitive load, whereas advanced learners can cope with indirect, unfocused WCF by drawing on their linguistic repertoire (Ellis, 2009; Kang & Han, 2015). Kang and Han's (2015) meta-analysis revealed that WCF effect sizes increased with proficiency, whereas for low-proficiency learners the impact was negative. Guénette (2007) argued that WCF effectiveness may depend on learners' cognitive developmental readiness and proficiency, advocating that proficiency be measured and treated as a moderating variable in WCF research — a recommendation that the studies above rarely follow, despite repeated calls for combined designs.

Language aptitude is a set of cognitive abilities, including phonetic coding, analytic ability, and rote memory, that support more efficient language acquisition (Dörnyei, 2005). Sheen (2007) reported a significant positive correlation between aptitude for language analysis and accuracy gains from direct and metalinguistic WCF, with metalinguistic WCF benefiting higher-aptitude learners more. Shintani and Ellis (2013) likewise found that higher-aptitude learners benefited more from both direct and metalinguistic WCF. Benson and DeKeyser (2019) reported contrasting results, with higher-aptitude learners benefiting more from direct WCF and lower-aptitude learners from metalinguistic feedback, likely due to differences in aptitude measures and feedback types. Michaud and McDonough (2025) found that language aptitude and working memory jointly

shaped how learners benefited from immediate versus delayed written CF during collaborative writing.

Most studies indicate that students generally prefer their teachers to correct their errors (Leki, 1991) and that preferences for certain feedback types influence uptake (Sinha & Nassaji, 2022). There remains a notable lack of research on the moderating role of L2 learners' attitudes in shaping the impact of WCF (Sinha & Nassaji, 2022). Storch and Wigglesworth (2010) found that the effectiveness of WCF was influenced by students' attitudes, beliefs, and objectives, and that these factors affected whether learning was retained over time. Rummel and Bitchener (2015) found that learners who received their preferred feedback type were more successful in correcting targeted errors. Roothooft and Bueno-Alastuey (2025) found that Spanish and American university students held different beliefs about WCF type and source, indicating that attitudes toward WCF vary by learning context.

To the best of the researchers' knowledge to date, no study has examined direct, indirect, and metalinguistic unfocused WCF within a single design while accounting for proficiency, aptitude, and attitudes, or has traced these effects into new-task performance after a delay using CAF measures. This is the gap the present study addresses.

Methods

This study employed a quasi-experimental design with four intact groups: direct CF (n = 40), indirect CF (n = 36), metalinguistic CF (n = 34), and a control group (n = 35) receiving no feedback. Intact groups were used because individual random assignment was not feasible within the institutional context; students were enrolled in pre-existing course sections. To verify baseline equivalence, a one-way ANOVA was conducted on pre-test CAF scores, Oxford Quick Placement Test (OPT) proficiency scores, and LLAMA_B and LLAMA_F scores across the four groups, revealing no statistically significant differences prior to treatment. Participants were assigned persuasive/argumentative essay topics from a curated list. Each participant wrote a practice essay, a pre-test essay, a post-test essay (a revised version of the pre-test), and a delayed post-test essay on a new topic. Each task comprised 10 minutes of planning and 30 minutes of writing, following established L2 writing research protocols (Skehan, 1998; Ellis et al., 2008) as sufficient to elicit meaningful written production while maintaining cross-session consistency.

Participants

The study was conducted at the American University of Sharjah, a co-educational private EMI university in the United Arab Emirates. All students were required to take an in-house placement test upon entering the University and were placed in either Academic Writing 101 or the lower-level Academic Writing 001 class. For this study, Academic Writing 001 students were selected because they had sufficient English writing proficiency to produce writing of the required length and level, yet their compositions still contained errors, allowing feedback to be provided. In total, 139 first-year students enrolled in Academic Writing 001 participated voluntarily in the study. Data were collected from four intact Academic Writing 001 classes. A convenience sampling approach was adopted, including all four available intact sections of Academic Writing 001 offered during the semester. The four intact groups were then randomly assigned at the class level to one of the four conditions (direct WCF, indirect WCF, metalinguistic WCF, or no feedback). Inclusion criteria required that participants be first-year students enrolled in Academic Writing 001 at the

time of data collection and that they consent to participate in the study. No participants were excluded from the final sample.

Ethical approval for the study was obtained from the University of Essex and, subsequently, from the American University of Sharjah, where the research was conducted. Prior to data collection, participants received consent forms and participation information sheets explaining the study and the tasks they would be required to complete. The first author also delivered a talk that elaborated on the study and addressed participants' questions. Participants were assured that their anonymity would be preserved in all study reports by removing or altering details that might reveal their identities and by assigning each student an identifier number, which was used to refer to their questionnaires, tests, or essays. Signed consent forms were then collected from all participants.

71 participants were female and 68 were male, with ages ranging from 18 to 20 years (mean age = 18.6). Most (115) were native Arabic speakers, while 24 were native Urdu speakers. Forty-five participants reported spending between 1 and 3 months in an English-speaking country, primarily the United Kingdom. Most participants (102) had attended English-medium schools, while 36 had attended Arabic-medium schools. Data on school type were not reported for one participant. The majority (121) had studied English for 12 years; 7 for 8 years; another 7 for 5 years; and 3 for 16 years. Only one participant did not provide this information.

Research Instruments

The attitude questionnaire, adapted from Leki (1991) and Saito (1994), comprised 22 Likert-scale items and three open-ended questions exploring students' attitudes towards corrective feedback. The OPT assessed grammar and vocabulary skills through 60 multiple-choice questions. The LLAMA language aptitude test (Meara, 2005) was selected for its language neutrality and widespread use in L2 research (Roehr-Brackin, 2020). LLAMA_B and LLAMA_F were selected for the present study because they tap the aptitude components most theoretically relevant to WCF processing: vocabulary learning ability (LLAMA_B), which underlies learners' capacity to acquire lexical corrections, and grammatical inferencing ability (LLAMA_F), which is central to processing indirect and metalinguistic feedback. The remaining sub-tests (LLAMA_D, LLAMA_E) assess phonological abilities with limited relevance to written performance and were therefore excluded. Participants' essays were assessed using five measures, as detailed in Table 1.

Table 1. *CAF measures used in the study*

Construct	Measure	Operationalization	Tool
Syntactic complexity	Complex nominals / clause	Number of complex nominals divided by number of clauses	LCA (Lu, 2010)
	Clauses per T-unit	Sentential syntactic complexity	LCA (Lu, 2010)
Accuracy	Error-free T-units (%)	Error-free T-units / total T-units; inter-rater agreement = 95%	Calculated by hand
Fluency	Total words	Total number of words in the text	LCA (Lu, 2010)
Lexical diversity	Vocd-D / MTLD	Vocd-D: predicted decline of TTR; MTLD: mean sequential word strings at TTR threshold = 0.72	Coh-Metrix

Note 1. LCA = L2 Syntactical Complexity Analyzer; MTLD = Measure of Textual Lexical Diversity

Accuracy, calculated by hand as the percentage of error-free T-units, was independently coded by a research assistant for 20% of the scripts to establish inter-rater consistency. The percentage agreement between the two markers was 95%. Fluency and syntactic complexity were derived using the L2 Syntactic Complexity Analyzer, an automated tool that produces identical output for a given text regardless of the rater and therefore did not require a second marker.

Research Procedures

During the initial sessions, students completed an online topic-selection questionnaire, a language-background questionnaire, and an attitude-towards-WCF questionnaire. They then completed two sub-tests of the LLAMA aptitude test (Meara, 2005)—sub-test B (vocabulary learning) and sub-test F (grammatical inferencing)—and the paper-based Oxford Quick Placement Test.

Following the pre-test essay (Session 6), treatment groups received WCF on their essays, while the control group received content feedback only. In post-feedback workshops lasting 30 minutes, students attempted to correct their essays: the direct feedback group could ask for clarification and received the correct forms; the indirect feedback group could ask only yes/no questions and had to find the correct forms independently; and the metalinguistic group used error codes and an accompanying legend to attempt self-corrections without receiving the correct forms. All four conditions were delivered by the same instructor to control for instructor-related variation in feedback delivery across groups.

In Session 8, students reviewed their essays and rewrote them from memory, without access to the original texts. This was a key methodological decision intended to prevent direct copying of corrections and to provide a more ecologically valid test of whether learning had occurred (cf. Truscott & Hsu, 2008; Van Beuningen et al., 2008). Four feedback rounds, each with a 30-minute post-feedback workshop, were completed over fourteen weeks. The delayed post-test, administered in Week 13 (five weeks after the final workshop), asked students to write a new argumentative essay on a new topic.

Data Analysis

Multivariate analysis of covariance (MANCOVA) was used to compare groups on multiple continuous dependent variables (CAF indices) while controlling for WCF type (independent variable) and attitudes, aptitude, and L2 proficiency (covariates). Separate MANCOVAs were conducted on pre-test to post-test gain scores and pre-test to delayed post-test gain scores. The dependent variables were continuous, linear, and normally distributed, with equal variance-covariance matrices across groups, thus meeting MANOVA assumptions (Field, 2013). Box's M and Levene's tests confirmed that additional MANCOVA assumptions were also met.

Results

RQ1: Effects of Feedback on Revised Tasks (Post-test)

Feedback type had a significant overall effect on writing performance gains in revised texts after controlling for individual difference variables [$F (15, 372) = 2.863$, $p < .0001$, Pillai's Trace $= .310$, partial $\eta^2 = .103$]. The medium effect size is consistent with those reported in recent WCF meta-analyses (Kang & Han, 2015; Sinha & Nassaji, 2022). The effects were specific rather than general: fluency ($F = 5.442$, $p < .001$, partial $\eta^2 = .115$) and lexical diversity ($F = 4.850$, $p = .003$, partial $\eta^2 = .104$) showed significant group differences with medium effect sizes, whereas

accuracy, syntactic complexity, and complex nominal density were unaffected by feedback condition (all ps > .14). None of the individual difference covariates reached statistical significance for any dependent variable. Full between-subjects results are presented in Table 2.

Table 2. *Tests of between-subject effects: Pre-test to post-test gain scores*

Source	Dependent Variable	Type III SS	df	Mean Square	F	Sig.	Partial η²
Group	Fluency gains	1138.262	3	379.421	5.442***	.001	.115
	Accuracy gains	349.773	3	116.591	1.675	.176	.038
	Complexity gains	118.748	3	39.583	0.563	.640	.013
	Lexical diversity gains	1622.498	3	540.833	4.850**	.003	.104
	Complex nominal gains	388.614	3	129.538	1.850	.141	.041
Proficiency (OPT)	Fluency gains	10.599	1	10.599	0.140	.709	.001
LLAMA_B	Fluency gains	2.662	1	2.662	0.035	.852	.000
LLAMA_F	Accuracy gains	130.000	1	130.000	1.632	.204	.013
Attitudes	Fluency gains	252.268	1	252.268	3.251	.074	.025

Note 2. ** p < .01; *** p < .001. Representative covariate rows shown per variable

The fluency findings revealed a counterintuitive pattern: more explicit feedback was associated with lower fluency gains. Indirect feedback outperformed direct feedback (mean difference = −5.821, p = .034), and—most strikingly—the no-feedback control group produced greater fluency gains than both the direct (mean difference = −7.463, p = .003) and metalinguistic conditions (mean difference = −5.993, p = .046). This suggests that explicit feedback may disrupt rather than support writing output during revision.

For lexical diversity, the pattern differed from that for fluency. No feedback type significantly outperformed the control group overall, but a meaningful distinction emerged among feedback types: both direct and metalinguistic feedback yielded greater gains in lexical diversity than indirect feedback (direct vs. indirect: mean difference = 7.602, p = .007; metalinguistic vs. indirect: mean difference = 8.156, p = .005). Indirect feedback, which offers no explicit lexical guidance, appears insufficient to scaffold learners' vocabulary development during revision. Full pairwise comparisons are presented in Table 3.

Table 3. *Pairwise comparisons for fluency and lexical diversity gains (Pre-test to Post-test)*

DV	(I) Group	(J) Group	Mean Diff. (I−J)	SE	Sig.	95% CI Lower	95% CI Upper
Fluency gains	Direct	Indirect	−5.821*	2.071	.034	−11.369	−0.274
		Control	−7.463*	2.116	.003	−13.133	−1.794
	Metalinguistic	Control	−5.993*	2.215	.046	−11.927	−0.059
	Control	Direct	7.463*	2.116	.003	1.794	13.133
		Metalinguistic	5.993*	2.215	.046	0.059	11.927
Lexical diversity gains	Direct	Indirect	7.602*	2.304	.007	1.430	13.774

ISSN: 2229-9327

	Metalinguistic	Indirect	8.156*	2.381	.005	1.779	14.534
	Indirect	Direct	−7.602*	2.304	.007	−13.774	−1.430
		Metalinguistic	−8.156*	2.381	.005	−14.534	−1.779

Note 3. Based on estimated marginal means. * $p < .05$. Bonferroni adjustment applied.

RQ2: Effects of Feedback on New Tasks (Delayed Post-test)

At the delayed post-test, when gains on new tasks were assessed, the advantage of feedback over no feedback disappeared entirely. The overall MANCOVA showed no significant group differences across the combined CAF measures [$F = 1.631$, $p = .064$, $\eta^2 = .062$]. Although the p-value approaches conventional significance, the small effect size suggests a marginal trend rather than a meaningful effect. At the level of individual CAF measures, a marginally significant group effect emerged for complex nominals per clause [$F (3, 126) = 2.774$, $p = .044$]; however, Bonferroni follow-up comparisons did not yield any statistically significant pairwise differences. None of the three feedback types produced measurable gains on new writing tasks relative to the control condition, and no feedback type outperformed any other.

RQ3: Influence of Individual Differences

None of the individual difference variables (L2 proficiency, language aptitude (LLAMA_B and LLAMA_F), or attitudes towards WCF) significantly moderated the effects of feedback on any CAF measure at either the post-test or delayed post-test stage. Regardless of their proficiency level, aptitude score, or attitudinal profile, learners responded similarly to each feedback type. In short, learners' characteristics did not determine how much they benefited from the type of feedback they received.

Discussion

RQ1: Effects of Feedback on Revised Writing Tasks

Feedback type had a selective rather than uniform effect on revised writing. Indirect feedback produced greater gains in fluency than direct or metalinguistic feedback, whereas direct and metalinguistic feedback produced greater gains in lexical diversity than indirect feedback. Neither accuracy nor syntactic complexity was significantly affected by feedback type.

The absence of an accuracy effect contrasts with Truscott and Hsu (2008) and Van Beuningen et al. (2008), who found that unfocused feedback improved accuracy during revision. A key methodological difference explains this discrepancy: in those studies, students retained their original texts alongside the feedback while revising, whereas in the present study, original texts were withheld before rewriting to prevent direct copying. This is a more demanding and ecologically valid test of learning, but it likely suppressed the short-term accuracy gains that a copying-supported design would produce.

The fluency finding is consistent with and extends models of limited attentional capacity in L2 performance. The reduction in fluency under direct and metalinguistic feedback likely reflects the explicit nature of these feedback types, which increased students' self-monitoring and cognitive load during revision. From a skill acquisition perspective (DeKeyser, 2015), this may reflect a temporary shift from proceduralised language production to controlled processing: when learners consciously attend to form during revision, attentional resources are diverted from meaning-focused production, reducing fluency. This aligns with Skehan's (1998) account of limited

attentional capacity, in which increased form monitoring can temporarily reduce production speed and output volume. Indirect feedback, being less explicit, may have been easier to engage selectively with or less disruptive to writing flow. An alternative interpretation is that the fluency decrease under explicit feedback reflects strategic adaptation rather than a writing deficit: learners receiving explicit feedback may deliberately prioritise correctness over quantity during revision, so lower fluency could indicate a considered shift in performance strategy rather than an inability to produce language.

The lexical diversity finding runs in the opposite direction and, to our knowledge, has not been reported alongside a fluency cost within the same unfocused WCF design. Explicit feedback appears to help learners identify and refine lexical choices, whereas indirect feedback, lacking such guidance, may lead learners to retain familiar but limited vocabulary. Because most prior unfocused WCF studies measured accuracy alone, they would not have been positioned to detect this fluency–lexical diversity trade-off. The present findings, therefore, indicate that gains in one CAF dimension can co-occur with losses in another within the same feedback condition. This means that accuracy alone is an incomplete proxy for feedback effectiveness, and that feedback type should be matched to the specific dimension of writing a teacher intends to develop.

These findings also have implications for classroom assessment. If fluency and lexical diversity move in opposite directions depending on feedback type, then assessment rubrics that weight these dimensions differently will present different pictures of the same learner's progress, depending on which unfocused WCF type their teacher happens to use. A fluency-weighted rubric may show a learner receiving direct or metalinguistic feedback as stagnating, while the same learner's lexical range is in fact improving, and the reverse holds for a learner receiving indirect feedback. This suggests that CAF-based classroom assessment following unfocused WCF should report gains across multiple dimensions separately rather than collapsing them into a single writing-quality score, since a composite score can mask a genuine trade-off as an absence of progress.

RQ2: Effects of Feedback on New Writing Tasks

At the delayed post-test, none of the three feedback types produced measurable gains on new writing tasks relative to the control condition, and no feedback type outperformed the others. This null result aligns with Truscott and Hsu (2008) and Van Beuningen et al. (2008), who likewise found limited accuracy gains on new texts. It contrasts with Bonilla López et al. (2018) and Nicolás-Conesa et al. (2019), who reported positive transfer effects. The difference is best explained by design rather than by feedback type: both of those studies combined WCF with additional instructional or motivational support, such as written language tasks or integration into graded coursework, whereas the present study tested unfocused WCF in isolation. Transfer to new tasks may therefore depend less on which unfocused WCF type is used than on whether feedback is embedded within a broader instructional scaffold, a distinction the existing comparisons do not isolate because feedback type and scaffolding are typically varied together rather than separately. This is consistent with transfer-appropriate processing accounts (Morris et al., 1977): WCF appears to support short-term error resolution rather than the deeper interlanguage restructuring needed for transfer, unless learning conditions are aligned with future production demands. Learners may successfully correct surface-level errors during revision without consolidating the underlying representations needed for transfer to new writing tasks.

RQ3: Influence of Individual Differences

None of the three individual-difference variables (L2 proficiency, language aptitude, or attitudes towards WCF) significantly moderated the effects of feedback on any CAF measure at either the post-test or the delayed post-test. This finding contrasts with Suzuki and DeKeyser (2017), who linked aptitude to feedback processing, and with the broader assumption in the individual-differences literature (e.g., Sheen, 2007; Kang & Han, 2015) that learner-internal variables shape feedback uptake. The absence of a proficiency effect may be attributable to the sample's relative homogeneity, as most participants were first-year university students with broadly similar English proficiency, or to the general nature of unfocused feedback, which does not target specific structures that learners at different proficiency levels are likely to process differently. The absence of an aptitude effect, despite the use of two aptitude measures, may reflect the same dilution caused by unfocused feedback. Unfocused feedback, unlike focused feedback targeting specific forms, does not create the inferential demands on which aptitude differences are most likely to bear. Self-reported positive attitudes did not predict performance gains, possibly because the attitudinal survey captured general preferences rather than task-specific engagement, or because attitudes alone are insufficient to sustain engagement without accompanying motivational support.

Taken together, these null results suggest that, in contexts involving unfocused WCF and general writing tasks, individual differences play a more limited role than the focused-feedback literature would predict. Individual differences reported for focused WCF may not generalise to unfocused WCF. The literature has assumed rather than tested this, since almost no prior study has examined the same ID variables across both feedback types within a single design. It is also possible that CAF measures are less sensitive to micro-level learner traits than measures targeting specific linguistic features or error types, which would support the broader argument that focused feedback may be more effective than unfocused feedback in contexts where individual differences are expected to play a role (Sheen et al., 2009; Ellis et al., 2008).

Conclusion

This study examined the effects of direct, indirect, and metalinguistic unfocused WCF on the CAF of L2 academic writing in a Gulf EMI context, and whether ID variables moderated learners' uptake of feedback. The study makes three novel contributions. To the researchers' knowledge, it is the first to compare all three unfocused WCF types simultaneously within a single design, using a comprehensive CAF framework. It incorporates both immediate and delayed post-tests to distinguish revision effects from transfer. Furthermore, it simultaneously examines three ID variables as potential moderators, thereby responding to calls in the literature for individual-difference research in WCF contexts (Li, 2015; Kang & Han, 2015).

In the short term, different feedback types significantly influenced fluency and lexical diversity, with medium effect sizes. Indirect feedback was most beneficial for fluency, whereas direct and metalinguistic feedback had an adverse effect on it. This pattern can be explained by the cognitive demands associated with more explicit feedback. Conversely, lexical diversity showed significant improvement for students receiving direct or metalinguistic feedback, but declined for those receiving indirect feedback, suggesting that more explicit feedback supports lexical development by helping learners identify and replace inappropriate lexical choices. No significant gains in accuracy or syntactic complexity were observed across any feedback condition in the short term.

In the delayed post-test, no statistically significant differences were observed across any CAF measures. This lack of long-term impact aligns with previous research indicating that unfocused WCF, particularly without sustained practice and support, has limited durability (Bitchener & Storch, 2016). These findings challenge the assumption that WCF alone yields comprehensive improvements in L2 writing and underscore the limitations of WCF used in isolation, without instructional integration or motivational scaffolding. ID variables had minimal moderating influence, which may reflect measurement limitations, sample homogeneity, or the possibility that moderating effects emerge over longer intervention periods.

Practical Recommendations for Teachers

Three recommendations for teachers emerge. First, feedback type should align with instructional goals. For example, indirect feedback is preferable when the goal is to promote fluency and writing flow, whereas direct or metalinguistic feedback is more appropriate when the aim is to expand lexical range and precision. Second, unfocused WCF should not be used in isolation, as the lack of significant transfer to new tasks underscores the need to embed feedback within structured revision cycles that include explicit instruction on how to process and respond to corrections. Third, teachers should be cautious about the time invested in unfocused WCF if learners lack the metacognitive strategies to engage meaningfully with it. Shorter, more targeted feedback episodes, integrated with revision support and goal-setting activities, are likely to yield more durable writing gains.

Limitations and Future Directions

Several limitations should be noted. The sample is restricted to first-year Gulf university students, limiting generalisability to other social, cultural, and educational contexts. The intervention period was relatively brief, and although a delayed post-test was included, the interval may have been insufficient to capture long-term retention or sustained writing development. The exclusive focus on unfocused WCF leaves unanswered questions about its comparative effectiveness relative to focused WCF under similar conditions.

Two additional limitations require acknowledgement. First, the study relied exclusively on quantitative CAF measures and did not collect qualitative data on learner engagement during feedback processing, such as think-aloud protocols, stimulated recall interviews, or written reflections. Without such data, it is not possible to determine how learners interpreted or acted on the feedback they received, limiting conclusions about the mechanisms underlying the observed effects. Second, CAF indices are global measures of written production and may not be sufficiently sensitive to detect micro-level linguistic changes that may occur even when overall scores remain stable. Future research should examine blended WCF approaches with extended treatment duration, integrate qualitative engagement data, and complement CAF analysis with targeted linguistic measures.

Funding: The authors received no funding for this study.
Acknowledgements: The authors are grateful to the participants and to the reviewers.
Conflicts of Interest: The authors declare no conflicts of interest.
Authenticity: The authors confirm that this manuscript is original and has not been published.
Artificial Intelligence Statement: The authors used Grammarly (an AI-based writing assistant) solely for copyediting and proofreading, including checks for grammar, spelling, and language

clarity. After using these tools, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

About the authors

Dr Laurence Craven is an Assistant Professor at the American University of Sharjah with a PhD in Applied Linguistics from the University of Essex. He has taught academic writing and communication in Japan, Spain, and the UAE. His research interests include corrective feedback, teaching writing and speaking, and AI and Computer-Assisted Language Learning (CALL). https://orcid.org/0000-0003-0739-2871

Dr Nasser Jabbari is an Assistant Professor in the School of Language, Literature, and Media at the University of Essex (UK), holding a PhD in Curriculum and Instruction from Texas A&M University (USA). With over two decades of teaching experience across diverse contexts, he explores second-language learning through digital games and social media and is currently researching the affordances and futures of GenAI in language education. https://orcid.org/0000-0001-5715-2265

References

Almohawes, M. (2025). Undergraduate EFL learners' preferences for three different types of written corrective feedback. *Frontiers in Education, 10,* Article 1532729. https://doi.org/10.3389/feduc.2025.1532729

Benson, S., & DeKeyser, R. M. (2019). Effects of written corrective feedback and language aptitude on verb tense accuracy. *Language Teaching Research, 23*(6), 702–726. https://doi.org/10.1177/1362168818770921

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing, 17*(2), 102–118. https://doi.org/10.1016/j.jslw.2007.11.004

Bitchener, J., & Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing*. Routledge. https://doi.org/10.4324/9780203832400

Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing, 19*(4), 207–217. https://doi.org/10.1016/j.jslw.2010.10.002

Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development*. Multilingual Matters.

Bonilla López, M., Van Steendam, E., Speelman, D., & Buyse, K. (2018). The differential effects of comprehensive feedback forms in the second language writing class. *Language Learning, 68*(3), 813-850. https://doi.org/10.1111/lang.12295

Chaudhary, A. (2022). The efficacy of continuous corrective feedback on writings of Saudi EFL students. *Arab World English Journal, 13*(4), 386–397. https://dx.doi.org/10.24093/awej/vol13no4.25

DeKeyser, R. M. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (2nd ed., pp. 94 –112). Lawrence Erlbaum Associates.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Lawrence Erlbaum Associates.

Ellis, R. (2008). *The study of second language acquisition*. Oxford University Press.

Ellis, R. (2009). Corrective feedback and teacher development. *L2 Journal, 1*(1), 3–18. https://doi.org/10.5070/l2.v1i1.9054

Ellis, R. (2010). Second language acquisition, teacher education and language pedagogy. *Language Teaching, 43*(2), 182–201. https://doi.org/10.1017/S0261444809990139

Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System, 36*(3), 353–371. https://doi.org/10.1016/j.system.2008.02.001

Ferris, D. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In F. Hyland & K. Hyland (Eds.), *Feedback in second language writing: contexts and issues* (pp. 81–104). Cambridge University Press. https://doi.org/10.1017/CBO9781139524742.007

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage.

Frear, D., & Chiu, Y. H. (2015). The effect of focused and unfocused indirect written corrective feedback on EFL learners' accuracy in new pieces of writing. *System, 53*, 24–34. https://doi.org/10.1016/j.system.2015.06.006

Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research, 29*(3), 311–343. https://doi.org/10.1177/0267658312461497

Guénette, D. (2007). Is feedback pedagogically correct?: Research design issues in studies of feedback on writing. *Journal of Second Language Writing, 16*(1), 40–53. https://doi.org/10.1016/j.jslw.2007.01.001

Hartshorn, K. J., & Pack, A. (2026). The effects of artificial intelligence-based dynamic written corrective feedback on second language writing and user sentiment. *RELC Journal, 57*(1). https://doi.org/10.1177/00336882251405498

Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *Modern Language Journal, 99*(1), 1–18. https://doi.org/10.1111/modl.12189

Karim, K., & Nassaji, H. (2020). The revision and transfer effects of direct and indirect comprehensive corrective feedback on ESL students' writing. *Language Teaching Research, 24*(4), 519–539. https://doi.org/10.1177/1362168818802469

Lalande, J. F., II. (1982). Reducing composition errors: An experiment. The Modern Language Journal, 66(2), 140–149. https://doi.org/10.1111/j.1540-4781.1982.tb06973.x

Leki, I. (1991). The preferences of ESL students for error correction in college-level writing classes. *Foreign Language Annals, 24*(3), 203–218. https://doi.org/10.1111/j.1944-9720.1991.tb00464.x

Li, S. (2015). The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics, 36*(3), 385–408. https://doi.org/10.1093/applin/amu054

Mahmood, R. Q. (2021). Kurdish EFL learners' perceptions towards written corrective feedback and its types: An investigative study. *Arab World English Journal, 12*(4), 103–117. https://doi.org/10.24093/awej/vol12no4.7

Mao, Z., Lee, I., & Li, S. (2024). Written corrective feedback in second language writing: A synthesis of naturalistic classroom studies. *Language Teaching, 57*(4), 449–477. https://doi.org/10.1017/S0261444823000393

Meara, P. (2005). *LLAMA language aptitude tests: The manual*. Lognostics.

Michaud, G., & McDonough, K. (2025). Task-based language teaching and the timing of written corrective feedback: The role of language aptitude and working memory. *System, 130*, 103622. https://doi.org/10.1016/j.system.2025.103622

Morris, P. E., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*(5), 519–533. https://doi.org/10.1016/S0022-5371(77)80016-9

Nicolás-Conesa, F., Manchón, R. M., & Cerezo, L. (2019). The effect of unfocused direct and indirect written corrective feedback on rewritten texts and new texts. *Modern Language Journal, 103*(4), 848–873. https://doi.org/10.1111/modl.12592

Oxford, R., & Shearin, J. (1994). Language learning motivation: Expanding the theoretical framework. *Modern Language Journal, 78*(1), 12–28. https://doi.org/10.2307/329249

Razmi, M. H., & Ghane, A. (2024). The impact of written corrective feedback on students' writing performance, self-efficacy, and anxiety. *Journal of Writing Research, 16*(2), 271–321. https://doi.org/10.17239/jowr-2024.16.02.04

Roehr-Brackin, K. (2020). LLAMA language aptitude tests. In S. Li, P. Hiver, & M. Papi (Eds.), *The Routledge handbook of second language acquisition and individual differences* (pp. 417–432). Routledge.

Roothooft, H., & Bueno-Alastuey, M. C. (2026). Students' beliefs about written corrective feedback in two foreign language contexts. *Language Awareness, 35*(2), 511–531. https://doi-org.uniessexlib.idm.oclc.org/10.1080/09658416.2025.2522107

Rummel, S., & Bitchener, J. (2015). The effectiveness of written corrective feedback and the impact Lao learners' beliefs have on uptake. *Australian Review of Applied Linguistics, 38*(1), 66-84. https://doi.org/10.1075/aral.38.1.04rum

Saito, H. (1994). Teachers' practices and students' preferences for feedback on second language writing: A case study of adult ESL learners. *TESL Canada Journal, 11*(2), 46-70. https://doi.org/10.18806/tesl.v11i2.633

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129–158. https://doi.org/10.1093/applin/11.2.129

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly, 41*(2), 255–283. https://doi.org/10.1002/j.1545-7249.2007.tb00059.x

Sheen, Y., Wright, D., & Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System, 37*(4), 556–569. https://doi.org/10.1016/j.system.2009.09.002

Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing, 22*(3), 286-306. https://doi.org/10.1016/j.jslw.2013.03.011

Sinha, T. S., & Nassaji, H. (2022). ESL learners' perception and its relationship with the efficacy of written corrective feedback. *International Journal of Applied Linguistics, 32*, 41–56. https://doi.org/10.1111/ijal.12378

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake, and retention of corrective feedback on writing. *Studies in Second Language Acquisition, 32*(2), 303–334. https://doi.org/10.1017/S0272263109990532

Suzuki, W., Nassaji, H., & Sato, K. (2019). The effects of feedback explicitness and type of target structure on accuracy in revision and new pieces of writing. *System, 81*, 135-145. https://doi.org/10.1016/j.system.2018.12.017

Suzuki, Y., & DeKeyser, R. (2017). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning, 67*(4), 747-790. https://doi.org/10.1111/lang.12241

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46*(2), 327–369. https://doi.org/10.1111/j.1467-1770.1996.tb01238.x

Truscott, J., & Hsu, A. Y. (2008). Error correction, revision, and learning. *Journal of Second Language Writing, 17*(4), 292–305. https://doi.org/10.1016/j.jslw.2008.05.003

Van Beuningen, C. (2010). Corrective feedback in L2 writing: Theoretical perspectives, empirical insights, and future directions. *International Journal of English Studies, 10*(2), 1-27.

Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning, 62*(1), 1–41. https://doi.org/10.1111/j.1467-9922.2011.00674.x

Van Beuningen, C., De Jong, N. H., & Kuiken, F. (2008). The effect of direct and indirect corrective feedback on L2 learners' written accuracy. *ITL International Journal of Applied Linguistics, 156*(1), 279–296. https://doi.org/10.2143/ITL.156.0.2034439

Westmacott, A. (2017). Direct vs. indirect written corrective feedback: Student perceptions. *Íkala, Revista de Lenguaje y Cultura, 22*(1), 17–32. https://doi.org/10.17533/udea.ikala.v22n01a02