



Empathic writing feedback: The role of teacher empathy in undergraduate students' writing development

 ZAHRA ZARGARAN  

University of Technology and Applied Sciences, Shinas, Oman

Abstract

Although feedback is widely regarded as one of the most influential variables in student learning, its effects on writing development remain highly variable and poorly understood, partly because the relational dimensions of feedback delivery have received limited empirical attention. This quasi-experimental, explanatory sequential mixed-methods study examined whether an Empathic Feedback Protocol (EFP), operationalising teacher empathy through three theoretically grounded dimensions (Seeing, Understanding, and Caring), produces superior outcomes in writing quality, self-efficacy, and feedback receptivity compared to conventional rubric-based feedback. Post Foundation undergraduate students in a Technical Writing course at a university in Oman were assigned to experimental or comparison conditions across an eight-week email writing programme. Writing quality was assessed using an institutional rubric; self-efficacy and receptivity were measured at three time points using validated scales; and student perceptions of feedback empathy were assessed using the EFRS-S, an instrument developed for this study. Qualitative data from semi-structured interviews and written reflections were analysed through reflexive thematic analysis and integrated with quantitative findings via joint display. Results indicated large condition effects on writing quality, with the greatest gains on higher-order criteria. Self-efficacy and feedback receptivity trajectories diverged significantly across conditions after the midpoint, with the experimental group sustaining growth while the comparison group plateaued. Hierarchical regression revealed a theoretically meaningful dissociation: Understanding was the dominant predictor of writing quality, while Caring predicted motivational outcomes. Qualitative themes corroborated these patterns and surfaced an emergent metacognitive transfer effect, with students spontaneously applying the EFP framework as a self-monitoring strategy during drafting. Findings suggest that teacher empathy, when systematically operationalised, is a consequential and teachable pedagogical variable with direct implications for feedback practice in EFL technical writing contexts.

Keywords: Teacher empathy, written corrective feedback, writing self-efficacy, feedback receptivity, EFL writing, email writing, mixed methods

Introduction

Teachers who are involved in writing, teaching, and assessment know the frustration of spending an hour on detailed comments only to watch students glance at their grade and slip the paper into a bag. Research confirms this is not an isolated experience. Meta-analytic reviews have consistently found that feedback produces measurable gains on average, yet the variation around that average is so wide that feedback also regularly fails to help, and occasionally makes things worse. Recent applied linguistics research also emphasizes the importance of aligning feedback with learners' sociocultural contexts (Hattie & Timperley, 2007; Mao & Crosthwaite, 2019; Wisniewski et al., 2020). In writing classrooms, the picture is particularly sobering; students misread feedback, ignore it, feel threatened by it, or engage only with the parts that explain why their grade fell short rather than with what might help them write better (Carless & Boud, 2018; Winstone et al., 2017). The research question that follows from all of this is not whether feedback is worth giving, but what distinguishes feedback that students actually use from feedback that they do not.

One answer that keeps surfacing in theoretical accounts but rarely gets tested empirically is teacher empathy. The idea has intuitive appeal; a teacher who genuinely tries to understand what a student was attempting when they wrote a particular sentence, and who communicates that understanding before offering correction, may produce a qualitatively different feedback experience than one who simply marks errors against a rubric. Decety and Jackson (2004) defined empathy as involving three interrelated processes: affective resonance with another's state, cognitive modelling of their perspective, and a maintained awareness of the distinction between self and other. Rogers (1969) and Noddings (1984) both made empathic understanding central to their accounts of good teaching, though from different theoretical starting points. A systematic review by Aldrup et al. (2022) found consistent associations between teacher empathy and positive student outcomes across a range of contexts. Yet, despite this theoretical and empirical backdrop, empathy has rarely been turned into something a teacher can deliberately do in writing feedback, something structured, teachable, and verifiable.

This study introduces the Empathic Feedback Protocol (EFP), which attempts to give teacher empathy a concrete and repeatable form in written feedback. The protocol centres on three dimensions. *Seeing* involves explicitly acknowledging what a student was trying to communicate and the effort visible in their writing, before any correction takes place. *Understanding* means identifying the probable underlying cause of a writing difficulty rather than just labelling the surface problem. *Caring* refers to writing in a way that communicates genuine investment in the student's progress, using language that looks forward rather than delivers a verdict, and that preserves the student's sense of themselves as a writer who is learning rather than one who is failing. These three dimensions draw on Decety and Jackson's (2004) neuroscientific model of empathy, the diagnostic strand of written corrective feedback research (Ellis, 2009; Ferris, 2011), and Noddings' (1984) ethics of care.

To investigate the integration of teacher empathy in writing feedback and its effect on writing quality of students, this study aims to answer four main research questions as follows:

1. Does empathy-informed feedback produce better email writing outcomes than conventional marking?
2. Does empathic feedback lead to different patterns of growth in writing self-efficacy and feedback receptivity over time?
3. Which of the three EFP dimensions—Seeing, Understanding, or Caring—best explains each type of outcome?

4. What is it actually like, from the student's perspective, to receive feedback structured this way?

Literature Review

Teacher Empathy and Instructional Effectiveness

Two quite different traditions have shaped how researchers and educators think about empathy in teaching. Rogers (1969) approached it as a psychological condition; when students feel genuinely understood, he argued, they become less defensive and more available for learning. This was not a minor addition to his theory of facilitative teaching but one of its three foundational elements, alongside congruence and unconditional positive regard. Noddings (1984) approached the question from moral philosophy rather than psychology. For Noddings, Caring, defined as attending to and responding to the full reality of the learner, is the ethical ground from which all good teaching grows. What matters here for the present study is Noddings' insistence that Caring is not merely a disposition but a practice—it shows up in specific acts of attention and response, not in how a teacher feels internally.

Aldrup et al.'s (2022) systematic review of 41 studies found positive links between teacher empathy and interaction quality, emotional support, and student outcomes across a range of educational settings. Their review also raised something worth taking seriously for feedback research: empathic teachers appeared better able to recognise when students were confused and to adjust their responses accordingly. That is, empathy was not just emotionally supportive, it was instructionally diagnostic. More directly relevant to the present study, Okonofua et al. (2016) showed in a randomised trial that an empathic stance in teacher communications reduced disciplinary incidents substantially, while Rabbani and Husain (2024) found that students who experienced feedback as harsh or dismissive showed measurable declines in motivation and academic self-concept. Caring feedback sustained engagement even when the corrective content was demanding.

Decety and Jackson's (2004) work on the neuroscience of empathy offers a useful framework for thinking about what empathic feedback actually requires of a teacher. Their model identifies three separable processes: affective sharing, which involves resonating with another person's emotional state; perspective-taking, which involves cognitively constructing a model of what the other person is experiencing; and self-other awareness, which is the capacity to remain grounded in one's own viewpoint while attending to another's. This distinction between affective empathy (resonating with another's emotional state) and cognitive empathy (accurately modelling another's perspective and mental processes) is theoretically consequential for feedback practice and is not merely terminological. Affective empathy unaccompanied by cognitive empathy risks yielding feedback that is diagnostically shallow, as the teacher resonates with the student's affective state without constructing an adequate explanatory model of the underlying source of difficulty. Feedback grounded in cognitive empathy alone risks being received as a clinical assessment rather than a meaningful response to a person. The student may process the correction accurately and still feel unseen. This condition can undermine engagement regardless of the feedback's instructional quality. Translated into the feedback context, the EFP's Seeing and Caring dimensions correspond most directly to the affective strand of empathy: Seeing acknowledges the student's felt purpose, while Caring communicates continued relational investment. Understanding, by contrast, operationalises cognitive empathy; it requires the teacher to construct an explanatory model of why a particular difficulty arose, grounded in where the student is developmentally and what their L1 background may be contributing. That three-part structure, resonating with effort, diagnosing difficulty, sustaining identity, maps

closely onto the EFP's Seeing, Understanding, and Caring dimensions; the affective/cognitive distinction maps onto the dissociation between motivational and learning outcomes that the present study's regression results subsequently confirm.

Written Corrective Feedback and EFL Writing Development

The written corrective feedback (WCF) literature has spent the better part of three decades arguing about whether grammatical correction helps L2 writers at all. Truscott's (1996) provocative claim that it does not, and may actively harm fluency development, generated a body of response work that has since produced a more qualified verdict. The consensus that has emerged, associated particularly with Ferris (2011) and Ellis (2009), is that focused feedback targeting specific features relevant to where a student actually is in their development tends to produce more durable gains than unfocused marking of everything that is wrong. Kang and Han's (2015) meta-analysis reinforced this, finding that the accuracy gains were larger still when feedback included metalinguistic explanation; that is, when teachers said something about why a particular form fails. That diagnostic move, explaining the nature of a difficulty rather than marking its location, is what the EFP's Understanding dimension is designed to formalise.

It is important, however, not to overstate this consensus. The WCF evidence base remains contested in meaningful ways. Truscott (2007) revisited his original position and maintained that the cumulative evidence still fails to demonstrate durable grammatical gains that transfer beyond the immediate correction context. Van Beuningen et al. (2012) found that comprehensive written feedback produced short-term accuracy improvements but that these did not fully hold when students revised independently, raising questions about whether observed gains reflect internalised learning or task-constrained performance. More recently, Xu and Harfitt (2019) found in a longitudinal study that students' emotional responses to correction and the quality of the feedback relationship were stronger predictors of revision behaviour than feedback type; it is a finding that the form-focused WCF literature has difficulty accounting for. These tensions suggest the field has prioritised the cognitive and linguistic dimensions of correction while leaving the affective and relational conditions that mediate feedback uptake comparatively underexplored. Hattie and Timperley's (2007) model makes a related point at a higher level of abstraction; feedback that engages with how students think about tasks and regulate their own learning consistently outperforms feedback that simply marks task-level errors. Lee's (2017) classroom research offered an illustration of what happens when feedback operates only at the evaluative level. Under pressure to avoid errors, English as a Foreign Language (EFL) students became progressively more conservative across drafts, producing technically safer but rhetorically flatter writing. They were, in effect, learning to write to what they thought the feedback would approve of rather than up toward what they were actually trying to say.

Email writing presents a specific challenge for EFL learners that purely accuracy-focused feedback struggles to address. What makes an email work or fail in a professional context is often not grammatical correctness but pragmatic appropriateness, the right level of formality, the right degree of directness, the right positioning of a request. Economidou-Kogetsidis (2018) and Hashemian and Farhang-Ju (2019) have both documented how mismatches in these areas are the primary source of communicative failure in L2 professional email writing. For Arabic-speaking learners specifically, the cross-linguistic influences in question are well-mapped; Al-Ali (2006) and Biesenbach-Lucas (2007) traced systematic patterns in how Arabic-English writers handle formality, directness, and epistemic stance. This matters for the EFP because it means the sources of pragmatic difficulty in this population are not random—they are identifiable and explicable, which makes Understanding-oriented feedback both possible and meaningful.

Writing Self-Efficacy and Feedback Receptivity in EFL Contexts

How confident a student feels about their writing shapes how they write. Students who believe they can manage the demands of a writing task are more likely to attempt complex language, take rhetorical risks, and revise substantively; those who doubt themselves tend to default to safe, minimal texts (Bandura, 1997; Bruning et al., 2013; Pajares, 2003). In EFL contexts, this relationship is particularly consequential because the gap between what students want to express and what they feel licensed to attempt in a second language is already wide (Golparvar & Khafi, 2021). Kara (2024), in a study of EFL university writing instruction, found that students given structured task-based writing opportunities showed significantly stronger writing competence gains, underscoring that the pedagogical framing of writing tasks shapes developmental outcomes as much as the correctness of individual responses. Teacher feedback occupies a central place in how self-efficacy beliefs form and change. In this vein, Bandura (1997) identified verbal persuasion from significant others, which in writing classrooms means the teacher, as one of the four primary sources of efficacy beliefs. In a meta-analysis, Ruegg (2018) and Lee and Evans (2019) both found that the quality and perceived usefulness of feedback mediated how much self-efficacy grew over a semester. Highly relevant to the present study, Busse et al. (2023) found that self-efficacy responded more strongly to feedback in the later stages of an intervention than at the start, suggesting that relational support accumulates in its effects rather than delivering a fixed boost at any single moment.

A complementary strand of research has shifted attention from what teachers write, to how students receive it. Feedback engagement is also socially mediated and shaped by emotional responses (Han & Hyland, 2019). Winstone et al.'s (2017) review of feedback intervention studies found that whether students actively sought out, interpreted, and acted on feedback—what they called proactive recipience—was at least as important a determinant of learning gains as the quality of the feedback itself. Carless and Boud (2018) developed a related framework under the label of feedback literacy, which includes, among other capacities, the ability to manage the emotional experience of receiving critical commentary. This is not a trivial skill, students who experience feedback as a threat to their identity as writers tend to focus their energy on defending that identity rather than on processing instructional content. Lipnevich et al. (2016) added an important nuance here by demonstrating that receptivity is not a stable personality trait. It responds to context—specifically, to the quality of the feedback relationship and the degree of safety a student feels within it. In Post Foundation EFL settings, where students are negotiating both a new genre and a professional identity in a second language, that sense of safety may be more fragile, and therefore more consequential, than in other contexts.

Several gaps in the existing literature motivated this study. Hyland and Hyland (2006) and Lipnevich and Smith (2009) have documented the kinds of relational moves skilled teachers make in feedback, but neither study attempted to formalise those moves into a protocol whose presence or absence could be verified and compared. The question of whether structured empathic feedback produces better outcomes than sole use of standard rubric-based marking therefore remains empirically open. Students know the difference between feedback that feels personal and feedback that feels like a grade. Fan et al. (2024), investigating English for Academic Purposes (EAP) teacher feedback on EFL disciplinary writing, confirmed this empirically: students who experienced feedback as relational and supportive engaged with it more deeply than those who experienced it as purely evaluative. Their findings affirm the rationale underlying the present study, while leaving open the question of whether structuring empathy explicitly as a feedback protocol produces still stronger effects. Similarly, Li and Xu (2024) found that the degree to which students perceived feedback as emotionally safe and personally directed was a significant predictor of their willingness to revise substantively. These studies

confirm the importance of the empathic dimensions of feedback, yet neither provides a replicable protocol for how such empathy can be systematically built into teacher feedback practice. A second gap concerns specificity; the existing research treats teacher empathy as a unified construct. Whether different components of empathy drive different kinds of outcomes, writing improvement on the one hand, motivational engagement on the other, has not been examined. A third gap is geographical and contextual. Research on empathy in feedback has been conducted almost entirely in Western and East Asian contexts. Arabic-speaking EFL learners in Gulf higher education, where both the language-learning trajectory and the social meaning of feedback carry distinctive features, have been largely absent from this literature.

Method

Research Design

The study used a quasi-experimental design with an explanatory sequential mixed-methods structure (Creswell & Plano Clark, 2018). Four intact class sections were assigned to conditions: two received feedback delivered through the EFP and two received the instructor's conventional rubric-based feedback. Students had been enrolled in these sections before the study began, so random assignment was not possible; statistical checks were run before any analysis to confirm the groups were comparable at baseline (Shadish et al., 2002). The quantitative data were collected across a pretest–posttest framework, and the qualitative strand followed afterwards, using semi-structured interviews and written reflections to explore and explain what the numbers showed. The two strands were then integrated through a joint display analysis (Fetters et al., 2013).

Participants

The study involved undergraduate students enrolled in a Technical Writing course at a public university in Oman across four sections. All were Arabic-speaking Post Foundation students at an intermediate level of English, aged mainly between 18 and 24. The participants were 117 students enrolled in four intact Technical Writing classes, with approximately 30 students in each class. All four sections were taught by the same instructor so that any differences between conditions could not be attributed to who was teaching. Ethical approval was obtained from the institutional research ethics committee before recruitment began, and students were told clearly that choosing not to participate would have no effect on their grades or standing.

The Empathic Feedback Protocol

According to the EFP, before marking anything as wrong, the researcher was required to acknowledge what the student had been trying to do and to name any effort visible in the draft; this was the Seeing dimension. Then the researcher needed to offer some account of why a particular difficulty had likely arisen, grounded in where the student was in their development and what their L1 Arabic background might be contributing, rather than simply identifying the surface error; this was Understanding. Finally, every feedback response needed to include language that looked forward rather than backward, communicating that the researcher expected the student to grow; this was Caring. Each piece of feedback for students in the experimental condition was required to contain at least one comment of each type.

Whether the feedback provided actually met the EFP criteria was checked by an independent rater who was not told which condition each sample came from. The rater coded all experimental feedback samples using the Feedback Empathy Coding Scheme, and the

agreement rate was high. A potential confound also needed to be addressed; if the experimental feedback was simply longer than the comparison feedback, any observed differences in outcomes might reflect quantity rather than quality. Word counts across both conditions were therefore compared and the difference was negligible and non-significant. Students in the comparison condition received the standard rubric-based written feedback that the researcher uses in all courses—this was a genuine professional baseline, not a deliberately weakened control.

Instruments

Writing Quality

Writing quality was assessed using the institutional email writing rubric in a university in Oman, which covers Task Achievement, Organisation, Grammar, Vocabulary, and Structural Elements. The rubric was applied at both the start and end of the eight-week period to emails of both types practised in the course, composing and replying. All samples were scored independently by a second rater with expertise in writing assessment who did not know which condition any student was in, and agreement between the two sets of scores was strong at both time points.

Writing Self-Efficacy Scale (WSES)

Writing self-efficacy was measured using the WSES (Bruning et al., 2013), a validated 15-item scale covering Ideation, Writing Conventions, and Self-Regulation on a five-point response scale. Internal consistency was good across subscales and time points.

Feedback Receptivity Scale (FRS)

Feedback receptivity was measured using an adapted version of the Proactive Recipience of Feedback Scale (Winstone et al., 2017). The adaptation retained the core structure of the original while fitting the specific course context.

Empathic Feedback Rating Scale—Student Form (EFRS-S)

Because no existing instrument measured students' perceptions of the three EFP dimensions specifically, a new scale was developed for this study, the Empathic Feedback Rating Scale, Student Form (EFRS-S). The development process followed established scale construction procedures (DeVellis, 2016). An initial item pool of 36 candidate items was generated deductively from the theoretical definitions of each EFP dimension and reviewed by three experienced EFL writing instructors and two applied linguists for relevance, clarity, and content representativeness. Items were refined iteratively based on reviewer feedback, and the pool was reduced to 24 items for pilot testing with a separate cohort of 40 students not included in the main study. Pilot data were examined for item-total correlations, response distribution, and evidence of ceiling or floor effects; nine items were removed or substantially reworded. The final instrument comprises 15 items distributed across three subscales corresponding to Seeing, Understanding, and Caring, each rated on a five-point response scale. Confirmatory factor analysis on the main study data confirmed that the three-factor structure fit the data well, and the reliability and convergent validity indicators all reached or exceeded recommended thresholds (Fornell & Larcker, 1981). This is the first time the EFRS-S has been used beyond its development context, so findings from analyses involving it should be treated as provisional until the scale is validated in independent samples and with different populations.

Procedure

In the first week, a research assistant administered the writing pretest and the two baseline surveys. The eight-week course then followed its normal schedule, with students completing five email writing assignments across Weeks 2 through 7. The two motivation surveys were re-administered at the midpoint and again at the end. A written reflection task and a final round of surveys followed the posttest writing task in Week 8. After the course ended, the researcher conducted audio-recorded interviews with 15 students from the experimental sections, chosen to represent a range of writing proficiency levels within each section. Transcripts were checked for accuracy and then five participants reviewed the thematic interpretations to make sure the analysis reflected their actual experiences.

Data Analysis

Before running any comparisons, the two groups were checked for pre-existing differences on all three baseline measures; none were found. Writing quality was then compared between conditions using ANCOVA, with pretest score included as a covariate. The self-efficacy and receptivity trajectories were examined through mixed-design ANOVA, which allowed the researcher to look at how both groups changed across three time points and whether those change patterns differed between conditions. To identify which EFP dimensions drove which outcomes, a series of hierarchical regression models were run, entering the baseline scores first and then the three EFRS-S subscales as predictors. Effect sizes are reported throughout. A clarification regarding the sample size for the regression analyses is warranted here. Although the EFRS-S was designed to capture student perceptions of the three EFP dimensions, the instrument was administered to all 117 participants rather than only to the experimental group. This decision was deliberate; students in the comparison condition were asked to rate the feedback they actually received (rubric-based feedback) on the same perceptual dimensions, which allowed the regression to examine whether perceived feedback characteristics, regardless of condition, predicted outcomes across the full sample. This approach treats perceived empathic quality as a continuous predictor rather than a binary condition variable, and it is consistent with how perception-based scales have been used in comparative feedback research (Lipnevich et al., 2016). That said, this approach is not without limitations; the EFRS-S items were developed with EFP feedback specifically in mind, and their conceptual applicability to rubric-based feedback is less straightforward. The interview and reflection data were analysed following the reflexive thematic analysis approach described by Braun and Clarke (2019), and the two strands were brought together through joint display (Fetters et al., 2013).

Results

Preliminary Analyses

Before examining condition differences, the two groups were checked for pre-existing differences on all outcome measures. As shown in Table 1, the gaps between conditions on all three baseline measures were negligibly small and statistically non-significant, confirming that the groups entered the study on essentially identical footing. These results also provided reassurance regarding the homogeneity of variance assumption underlying the planned parametric analyses. Scale reliability was strong across all instruments and time points, and the factor structure of the EFRS-S held up under confirmatory analysis.

Table 1. *Baseline equivalence: Descriptive statistics and independent-samples t-test results*

Measure	Exp. <i>M</i> (<i>SD</i>)	Comp. <i>M</i> (<i>SD</i>)	<i>t</i>	<i>df</i>	<i>p</i>	<i>d</i>
Writing Quality (pretest)	6.62 (0.87)	6.57 (0.88)	0.31	115	.758	.057
WSES (baseline)	2.91 (0.53)	2.88 (0.56)	0.30	115	.767	.055
FRS (baseline)	3.04 (0.48)	3.01 (0.51)	0.33	115	.744	.061

Note. Exp. = Experimental (*n* = 58); Comp. = Comparison (*n* = 59). WSES = Writing Self-Efficacy Scale; FRS = Feedback Receptivity Scale. All *p*-values two-tailed. *d* = Cohen’s *d*. No significant differences detected (all *p* > .05).

Table 2 shows how the groups performed on the writing rubric before and after the intervention. Students who received EFP feedback improved substantially more than those in the comparison condition, and this difference remained large and statistically significant after accounting for where each group started. The adjusted gap at posttest represented a meaningful practical difference in overall email writing quality.

Breaking the results down by subscale reveals a pattern worth noting. The biggest gains were on Task Achievement and Organisation, the two criteria most directly concerned with whether the student’s email writing performance demonstrated the intended communicative purpose and was coherently structured. The EFP also appeared to help with vocabulary and grammar, though somewhat less dramatically. There was no meaningful difference between conditions on Structural Elements, the formatting conventions of email structure, which makes sense, a rubric tells students what format is expected, and that is sufficient.

Table 2. *Writing quality by condition: Pretest, Posttest, and ANCOVA (Rubric /10)*

Subscale	Exp. Pre <i>M</i> (<i>SD</i>)	Exp. Post <i>M</i> (<i>SD</i>)	Comp. Pre <i>M</i> (<i>SD</i>)	Comp. Post <i>M</i> (<i>SD</i>)	<i>F</i> (1,114)	ηp^2
Task Achievement (/4)	2.71 (0.61)	3.74 (0.44)	2.69 (0.58)	3.12 (0.52)	44.82***	.28
Organisation (/2)	1.31 (0.38)	1.81 (0.29)	1.29 (0.41)	1.48 (0.34)	38.64***	.25
Vocabulary (/1.5)	0.91 (0.28)	1.24 (0.22)	0.92 (0.31)	1.04 (0.27)	17.23***	.13
Grammar (/1.5)	0.88 (0.31)	1.14 (0.26)	0.87 (0.29)	1.02 (0.28)	12.08**	.09
Structural Elements (/1)	0.81 (0.27)	0.88 (0.24)	0.80 (0.29)	0.86 (0.26)	1.94	.02
Total Score (/10)	6.62 (0.87)	8.81 (0.67)	6.57 (0.88)	7.52 (0.78)	49.17***	.30

Note. The writing rubric yields a total score out of 10, comprising five analytic criteria: Task Achievement (/4), Organization (/2), Grammar (/1.5), Vocabulary (/1.5), and Structural Elements (/1). Total scores are the sum of the five subscale scores. Means and *SD*s in all cells are based on observed scores. Post-test Cohen’s *d* values were calculated using pooled observed posttest standard deviations prior to ANCOVA adjustment. *** *p* < .001, ** *p* < .01.

The self-efficacy and receptivity findings illustrated a similar and theoretically coherent pattern. Both groups showed some improvement between the start and the midpoint. But, then their trajectories diverged, the experimental group continued to grow from midpoint to posttest while the comparison group’s gains levelled off. By the end of the study the gap between conditions on both measures was large. Table 3 presents the full means across all three time points.

Table 3. *WSES and FRS means (SD) by condition and time point*

Measure / Condition	T1 Pre M (SD)	T2 Mid M (SD)	T3 Post M (SD)	F interact.	np ²	Post d
WSES — Experimental	2.91 (0.53)	3.29 (0.48)	3.80 (0.44)	22.47***	.16	1.04
WSES — Comparison	2.88 (0.56)	3.14 (0.51)	3.26 (0.49)			
FRS — Experimental	3.04 (0.48)	3.41 (0.44)	3.88 (0.41)	18.34***	.14	1.02
FRS — Comparison	3.01 (0.51)	3.22 (0.49)	3.31 (0.47)			

Confirmatory factor analysis (CFA) was conducted on the EFRS-S to verify its hypothesized three-factor structure (Seeing, Understanding, and Caring) using the full study sample ($N = 117$). The model showed acceptable to good fit to the data. As shown in Table 4, all fit indices met or exceeded recommended thresholds. All factor loadings were statistically significant and ranged from .68 to .89. The subscales also showed strong convergent validity and internal consistency, with Average Variance Extracted (AVE) values of .62 (Seeing), .68 (Understanding), and .71 (Caring), and composite reliability (CR) coefficients of .84, .87, and .89, respectively. Inter-factor correlations ranged from .52 to .71, supporting the distinct but related nature of the three dimensions.

Table 4. *Confirmatory factor analysis for the empathic feedback rating scale*

Fit Index	Value	Recommended Cut-off	Interpretation
χ^2 (df)	142.36 (87)	—	—
χ^2/df	1.64	< 3.00	Good fit
CFI	.96	≥ .95	Good fit
TLI	.95	≥ .95	Good fit
RMSEA [90% CI]	.062 [.042, .080]	≤ .08	Acceptable fit
SRMR	.051	≤ .08	Good fit

The regression models asked which of the three dimensions was actually doing the most work. The answer differed depending on which outcome was being predicted. For writing quality, Understanding was the strongest predictor, followed by Seeing; Caring did not add significantly to writing quality once the other dimensions were accounted for. For both self-efficacy and receptivity, the pattern reversed, Caring was clearly dominant, with Understanding playing a secondary supporting role, and Seeing not contributing independently. Table 5 reports the full

regression results. The split is theoretically coherent; Understanding improves what students write because it gives them better diagnostic information, while Caring improves how students feel about feedback and about themselves as writers.

Table 5. Hierarchical regression block 2: EFRS-S subscales predicting posttest outcomes (N = 117)

Predictor	Writing Quality β	WSES β	FRS β	ΔR^2	p
Seeing	.26**	.08	.09	.17	< .001
Understanding	.37***	.23**	.20*	.21	< .001
Caring	.10	.44***	.45***		

Analysis of the interviews and written reflections produced four themes. The first, *being read rather than graded*, described something students noticed immediately about the EFP feedback—it treated their writing as a communicative act with a purpose behind it, not just a performance to be scored. Many contrasted this with how feedback had felt in previous courses. The second theme, *understanding why not just what*, captured students’ responses to root-cause explanation in feedback. Several described how knowing where a problem came from, rather than just that it existed, allowed them to catch similar issues themselves in later assignments. The third theme, *feedback as relational communication*, described a shift in how students approached the feedback itself; the caring-oriented framing made them less defensive and more willing to read comments carefully even when those comments were critical. The fourth theme emerged unexpectedly and may be the most interesting; several students described having begun to use the Seeing, Understanding, Caring structure as a way of reading their own drafts before submitting them. This was not something the study had planned to investigate; it appeared spontaneously in multiple accounts.

Table 6 presents the joint display mapping between the quantitative and qualitative findings. The patterns converged closely across all three dimensions. One additional insight came from the qualitative data alone—students who did not feel their work had been seen as purposeful described the caring-oriented comments in their feedback as hollow, something they recognised as well-intentioned but could not quite believe. This suggests Seeing functions less as an independent predictor and more as the precondition that makes the other two dimensions credible. When it is absent, Caring loses its authenticity.

Table 6. Joint display: Mixed-methods integration by EFP dimension

EFP Dimension	Quantitative Finding	Qualitative Theme	Integration Type	Inference
Seeing	Largest effects on Task Achievement ($\eta p^2 = .28$) and Organization ($\eta p^2 = .25$); $\beta = .26$ for writing quality	Theme 1: Being Read Rather Than Graded — intent acknowledgement reduces communicative risk-aversion	Convergence	Seeing validates writerly intent, enabling higher-order risk-taking

(Continued)

EFP Dimension	Quantitative Finding	Qualitative Theme	Integration Type	Inference
Understanding	Strongest predictor of writing quality ($\beta = .37$); also significant for WSES ($\beta = .23$) and FRS ($\beta = .20$)	Theme 2: Understanding Why Not Just What — root-cause diagnosis enables transfer across assignments	Convergence	Understanding converts individual corrections into transferable diagnostic frameworks
Caring	Dominant predictor of WSES ($\beta = .44$) and FRS ($\beta = .45$); non-significant for writing quality ($\beta = .10$)	Theme 3: Feedback as Relational Communication — relational framing reduces defensive orientation	Convergence	Caring operates through motivational-relational pathway, sustaining engagement across feedback cycles
All three	Seeing: non-significant for WSES/FRS in isolation; acts as prerequisite condition	Theme 4: Feedback as Model for Self-Assessment — students spontaneously internalized EFP as self-monitoring heuristic	Elaboration	Seeing authenticates the encounter; without it, Caring cannot be received as genuine

Discussion

The findings reported above support a fairly specific claim: teacher empathy, when given a structured form in written feedback, does more than criticizing students’ writing performance. It changes how they write, and it changes how they feel about their capacity to write. These are not the same thing, and the regression results suggest they are driven by different components of the empathy framework. That dissociation, Understanding predicting skill, Caring predicting motivation, is the finding that is most worth dwelling on, because it implies that the existing tendency to treat teacher empathy as a unified disposition is probably naive. The discussion that follows works through the key results in turn, considering what they add to the existing literature and what they leave open.

Empathy-Informed Feedback and Email Writing Quality

The large overall effect on writing quality makes sense in light of what the EFP is actually doing. Hattie and Timperley (2007) identified feedback that operates at the level of how students approach tasks and monitor their own learning as substantially more powerful than feedback that marks surface errors. The EFP’s Seeing and Understanding dimensions are both designed to work at that higher level, Seeing by positioning the student as a writer with genuine communicative intentions, Understanding by explaining rather than merely identifying difficulty. Neither of these is what a rubric does on its own. A rubric tells students how their work was evaluated; the EFP adds an account of how the teacher read it and what the student might usefully think about next. This is broadly consistent with Bitchener and Storch’s (2016) argument

that L2 feedback needs to engage with the communicative and rhetorical dimensions of writing, not just its surface accuracy.

Task Achievement and Organisation, the criteria where a student's communicative choices are most visible, showed the largest gains, while Structural Elements, which covers formatting conventions taught directly in class, showed no meaningful difference between conditions. This tracks well with Lee's (2017) finding that purely evaluative feedback pushes students toward safe, formulaic choices. The EFP's Seeing dimension directly addresses that dynamic by validating the student's communicative intentions before the correction begins, giving students a reason to keep trying to say what they mean rather than retreating to the least risky available formulation. Format conventions, by contrast, are just rules, and a rubric shows students whether they followed them. There is no developmental gap there for empathic feedback to fill.

Before drawing stronger conclusions from these patterns, three alternative explanations deserve serious consideration. The first is a novelty effect. Students in the experimental condition received a qualitatively different kind of feedback than they had encountered before, and it is possible that the gains observed reflect heightened engagement produced by novelty and increased attention rather than anything specifically empathic about the protocol. This account is difficult to fully rule out in any single-cohort study, and it would predict that the gains observed for the experimental group would erode in subsequent semesters as the feedback approach became familiar. It is a hypothesis that longitudinal follow-up could test. The second alternative is teacher enthusiasm. The researcher, as the EFP developer, may have delivered the experimental feedback with greater investment and attentiveness than the comparison feedback, irrespective of whether that difference was visible in word count or protocol compliance checks. Research on teacher expectancy effects documents that differential enthusiasm can produce measurable differences in student performance even when instructional content is held constant (Jussim & Harber, 2005; Rosenthal & Jacobson, 1968). The independent rater confirmed that the EFP criteria were met, but rater checks cannot detect the quality of teacher engagement that students may have sensed through other means. The third alternative concerns student perception bias. Students who received more personalised, elaborate feedback may have inferred that their instructor cared more about their progress and responded accordingly, not because the EFP dimensions themselves drove learning, but because perceived differential investment motivated greater effort. These three explanations are not mutually exclusive, and they are not easily separable from the causal account this study favours. Acknowledging them is not to undermine the findings but to be precise about what this design can and cannot establish. A dismantling study that delivers each EFP dimension in isolation, by a researcher not invested in the outcome, to groups unaware of condition differences, would go considerably further toward isolating the mechanisms involved.

Self-Efficacy and Feedback Receptivity Trajectories

The motivational trajectory results are worth reading carefully. Both groups improved in self-efficacy and receptivity between the start and the midpoint of the course, which is not surprising; any structured feedback intervention tends to boost engagement early on (Nicol & Macfarlane-Dick, 2006). What distinguished the conditions was what happened after that. The comparison group plateaued. The experimental group kept going. By the end of the eight weeks, the gap between conditions was large on both measures. This cumulative pattern matches what Busse et al. (2023) found in their longitudinal work, writing self-efficacy is more sensitive to feedback in the later stages of an intervention, as the relational experience of feeling supported accumulates across encounters. A single caring comment probably does not change how a student feels about themselves as a writer. Five or six, over several weeks, apparently can.

Bandura (1997) identified verbal persuasion from significant others as one of the four main sources of self-efficacy beliefs. In a writing course, the teacher is the primary source of that persuasion, and what they say, and how, determines whether the persuasion builds confidence or erodes it (Zimmerman & Bandura, 1994). The EFP's Caring dimension is designed to make that persuasion explicitly positive and developmental rather than evaluative, which appears to have sustained efficacy growth beyond the initial boost both groups showed. The receptivity finding adds a complementary mechanism—Winstone et al. (2017) showed that students' sense of relational safety shapes how openly they engage with feedback. Caring seems to operate directly on that sense of safety, making students more willing to read comments closely and act on them, rather than reading defensively or avoiding the feedback altogether.

The Differential Predictive Role of the Three EFP Dimensions

The regression results suggest that Understanding and Caring are not simply two expressions of the same underlying empathy. They predict different things through different mechanisms. Understanding is associated with higher writing quality because it provides students with more actionable information. For example, when a teacher explains that a particular pragmatic misjudgement in an email likely reflects Arabic–English transfer patterns rather than carelessness, the student has something specific and transferable to work with (Ferris, 2011; Kang & Han, 2015). The cross-linguistic context matters here. For Arabic-speaking learners, the pragmatic sources of email writing difficulty are not idiosyncratic, they are systematic and identifiable (Al-Ali, 2006; Biesenbach-Lucas, 2007). Understanding-oriented feedback can name them, which makes the correction instructionally durable rather than just locally corrective.

Caring works differently. It was not, on its own, directly associated with better writing quality. Rather, it was associated with how students encountered the feedback, whether they approached it defensively or openly, and whether they interpreted critical comments as threatening or useful. Lipnevich et al. (2016) showed that feedback receptivity responds to the quality of the relational context, and Noddings' (1984) philosophical account of caring provides the conceptual scaffolding; caring, for Noddings, is enacted in specific acts of attention and response that position the learner as a person whose growth matters, not just as a recipient of evaluation. The joint display analysis added one further nuance that the regression alone could not reveal; students who did not feel their work had been seen described the caring-framed comments as somehow not quite believable. The qualitative findings suggested that Seeing was closely associated with students perceiving Caring as genuine rather than performative. The three dimensions are not independent units; they form an ordered sequence in which Seeing was associated with students' receptivity to Caring, and Caring was, in turn, associated with the motivational conditions within which Understanding appeared to do its instructional work.

The Metacognitive Transfer Effect

The metacognitive transfer finding is the one that surprised the researcher most. Students in the experimental condition began using the Seeing–Understanding–Caring framework to read their own drafts, without being asked to. That is not a finding the study was designed to produce; it emerged from the interview and reflection data independently. Carless and Boud (2018) describe the highest-level feedback literacy as students becoming able to apply expert-reader criteria to their own work proactively rather than waiting for external feedback. If the EFP produces that capacity as an incidental outcome, it is not through explicit instruction in self-assessment but simply through repeated exposure to a particular way of being read. Pajares' (2003) work on self-efficacy adds a further dimension; mastery experiences built through successful self-assessment are a reliable source of efficacy growth, and they operate independently

of what the teacher writes. A student who has internalised the EFP as a self-monitoring tool may be building self-efficacy even between feedback cycles. Whether teaching the framework explicitly as a self-assessment protocol would produce stronger effects than this spontaneous internalisation is a question worth pursuing.

Implications for Practice and Future Research

The practical takeaway is relatively straightforward. The EFP does not ask teachers to abandon rubrics or to transform their entire feedback practice. It asks them to do three additional things, namely what the student was trying to achieve before correcting them and write in a way that looks forward rather than delivers a verdict. These are teachable habits, and professional development programmes could incorporate them without restructuring existing course designs. Given that format conventions showed no meaningful condition effect, instructors working with this population could reasonably concentrate the EFP effort on the higher-order criteria, Task Achievement, Organisation, and the accuracy dimensions where the protocol showed the greatest leverage.

Several research directions follow from the findings. The metacognitive transfer result raises an obvious next question; if students spontaneously adopted the EFP framework as a self-assessment tool, would explicitly teaching it as such produce still stronger or faster effects? A study that compares teacher-delivered EFP feedback with student-directed EFP self-assessment, or combines both, would be informative. The finding that Seeing appears to function as a prerequisite for the other dimensions rather than as an independent predictor deserves formal investigation through a dismantling design that systematically removes one dimension at a time. That kind of design would allow direct causal attribution and would settle the question of whether all three components are necessary or whether context-specific simplifications are defensible. Replication with different L1 groups and in different writing genres is needed before the framework can be considered generalisable.

Limitations

Five limitations are worth naming clearly. The design cannot support fully causal claims; without random assignment, the possibility that the two groups differed in ways the baseline measures did not capture cannot be entirely ruled out. The researcher's dual role as instructor and investigator is a more serious concern, not just methodologically, but epistemically. No set of procedural safeguards entirely removes the possibility that students sensed what the study expected of them, or that the thematic analysis was shaped by the researcher's investment in the protocol. The findings should be read as promising and hypothesis-generating rather than definitive. The single-institution sample limits how far the results can travel; a different institution, a different country, a different L1, may produce different patterns. The eight-week window is relatively short for tracking writing development; the gains observed here need to be followed up to see whether they persist. And the EFRS-S is a new instrument that has only ever been used in one study, which means all findings involving it must be treated with appropriate caution until the scale has been validated elsewhere.

Conclusion

Teacher empathy, given a concrete and structured form through the EFP, appears to make a genuine difference in both how Post Foundation EFL students write and how they feel about writing. The differences observed across eight weeks were large enough to matter practically, and the mixed-methods design gives some confidence that the quantitative patterns correspond to real experiences rather than statistical artefacts. What the study reveals is that treating teacher

empathy as a teachable practice rather than an innate disposition is worth pursuing, and that disaggregating empathy into components that do different instructional work is a more productive theoretical move than treating it as a unified variable. The finding that students began reading themselves through the EFP framework points toward a possibility that seems worth taking seriously. It considers that the best outcome of empathic feedback is not that students feel better about receiving it, but that they begin to write with it in mind.

References

- Al-Ali, M. N. (2006). Constructing the discipline: A register analysis of Arabic and English academic discourse in a Gulf university. *Journal of Applied Linguistics*, 3(2), 143–166.
- Aldrup, K., Carstensen, B., & Klusmann, U. (2022). Is empathy the key to effective teaching? A systematic review of its association with teacher and student outcomes. *Educational Psychology Review*, 34(3), 1177–1216. <https://doi.org/10.1007/s10648-021-09649-y>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.
- Biesenbach-Lucas, S. (2007). Making requests in email: Do cyber-consultations entail directness? Toward conventions in a new genre. *Language and Internet*, 4(1), 7.
- Bitchenor, J., & Storch, N. (2016). *Written corrective feedback for L2 development*. Multilingual Matters.
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4), 589–597. <https://doi.org/10.1080/2159676X.2019.1628806>
- Bruning, R., Dempsey, M., Kauffman, D. F., McKim, C., & Zumbrunn, S. (2013). Examining dimensions of self-efficacy for writing. *Journal of Educational Psychology*, 105(1), 25–38. <https://doi.org/10.1037/a0029692>
- Busse, V., Hennies, C., & Kreft, A. (2023). Writing self-efficacy in EFL secondary classrooms: A mixed-methods longitudinal study. *System*, 112. <https://doi.org/10.1016/j.system.2023.103001>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315–1325. <https://doi.org/10.1080/02602938.2018.1463354>
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.
- Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, 3(2), 71–100. <https://doi.org/10.1177/1534582304267187>
- DeVellis, R. F. (2016). *Scale development: Theory and applications* (4th ed.). SAGE.
- Economidou-Kogetsidis, M. (2018). Pragmatic failure, rating and critical comments in native speakers' evaluations of L2 emails. *Pragmatics and Society*, 9(1), 72–101. <https://doi.org/10.1075/ps.9.1.03eko>
- Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, 63(2), 97–107. <https://doi.org/10.1093/elt/ccn023>
- Ferris, D. R. (2011). *Treatment of error in second language student writing* (2nd ed.). University of Michigan Press.
- Fan, Y., Tan, S., & Lim, G. Y. W. (2024). EAP teacher feedback in the age of AI: Supporting first-year students in EFL disciplinary writing. *Australian Journal of Applied Linguistics*, 7(3), 1943. <https://doi.org/10.29140/ajal.v7n3.1943>
- Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—principles and practices. *Health Services Research*, 48(6), 2134–2156. <https://doi.org/10.1111/1475-6773.12117>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.2307/3151312>
- Golparvar, S. E., & Khafi, A. (2021). The role of L2 writing self-efficacy in integrated writing performance. *Assessing Writing*, 47, 100504. <https://doi.org/10.1016/j.asw.2020.100504>
- Hashemian, M., & Farhang-Ju, M. (2019). Applied linguistics faculty members' perceptions of (Im) politeness and (In)appropriateness of L2 learners' e-mail requests. *Teaching English as a Second Language Quarterly*, 38(1), 119–155. <http://doi.org/10.22099/jtls.2019.34578.2729>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>

- Han, Y., & Hyland, F. (2019). Academic emotions in written corrective feedback situations. *Journal of English for Academic Purposes*, 38, 1–13. <https://doi.org/10.1016/j.jeap.2018.12.003>
- Hyland, F., & Hyland, K. (2006). Feedback on second language students' writing. *Language Teaching*, 39(2), 83–101. <https://doi.org/10.1017/S0261444806003399>
- Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, 9(2), 131–155. https://doi.org/10.1207/s15327957pspr0902_3
- Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *The Modern Language Journal*, 99(1), 1–18. <https://doi.org/10.1111/modl.12189>
- Kara, S. (2024). Enhancing students' writing competence and 21st-century skills simultaneously via infographics. *Australian Journal of Applied Linguistics*, 7(3), 2044. <https://doi.org/10.29140/ajal.v7n3.2044>
- Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Springer.
- Lee, E., & Evans, M. (2019). Feedback dynamics and writing self-efficacy in peer feedback contexts. *Written Communication*, 36(3), 392–422. <https://doi.org/10.1177/0741088319865510>
- Li, M., & Xu, J. (2024). Emotional safety and revision behaviour in EFL writing: The mediating role of perceived feedback quality. *Journal of Second Language Writing*, 64, 101095. <https://doi.org/10.1016/j.jslw.2024.101095>
- Lipnevich, A. A., Berg, D. A. G., & Smith, J. K. (2016). Toward a model of student response to feedback. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 169–185). Routledge. <https://doi.org/10.4324/9781315749136-17>
- Lipnevich, A. A., & Smith, J. K. (2009). Effects of differential feedback on students' examination performance. *Journal of Experimental Psychology: Applied*, 15(4), 319–333. <https://doi.org/10.1037/a0016262>
- Mao, Z., & Crosthwaite, P. (2019). Investigating written corrective feedback: (Mis)alignment of teachers' beliefs and practice. *Journal of Second Language Writing*, 45, 46–60. <https://doi.org/10.1016/j.jslw.2019.05.004>
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218. <https://doi.org/10.1080/03075070600572090>
- Noddings, N. (1984). *Caring: A relational approach to ethics and moral education*. University of California Press.
- Okonofua, J. A., Paunesku, D., & Walton, G. M. (2016). Brief intervention to encourage empathic discipline cuts suspension rates in half among adolescents. *Proceedings of the National Academy of Sciences*, 113(19), 5221–5226. <https://doi.org/10.1073/pnas.1523698113>
- Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Reading & Writing Quarterly*, 19(2), 139–158. <https://doi.org/10.1080/10573560308222>
- Rabbani, L. M., & Husain, S. H. (2024). Fostering student engagement with criticism feedback: Importance, contrasting perspectives and key provisions. *Frontiers in Education*, 9, 1344997. <https://doi.org/10.3389/feduc.2024.1344997>
- Rogers, C. R. (1969). *Freedom to learn: A view of what education might become*. Merrill.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. Holt, Rinehart & Winston.
- Ruegg, R. (2018). The effect of peer and teacher feedback on changes in EFL students' writing self-efficacy. *The Language Learning Journal*, 46(2), 87–102. <https://doi.org/10.1080/09571736.2014.958190>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- 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. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, 16(4), 255–272. <https://doi.org/10.1016/j.jslw.2007.06.003>
- 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>

- Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational Psychologist*, 52(1), 17–37. <https://doi.org/10.1080/00461520.2016.1207538>
- Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.03087>
- Xu, J., & Harfitt, G. J. (2019). The role of the teacher–student relationship in written corrective feedback: A mixed-methods study. *TESOL Quarterly*, 53(4), 984–1010. <https://doi.org/10.1002/tesq.512>
- Zimmerman, B. J., & Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *American Educational Research Journal*, 31(4), 845–862. <https://doi.org/10.3102/00028312031004845>