

Measuring Cognitive Engagement in AI-Supported Writing: Introducing the Revision Responsiveness Index (RRI) as a Process-Based Assessment Model

Traci C. Eshelman, PhD

Liberty University

teshelman@liberty.edu

ORCID: <https://orcid.org/0000-0003-0234-8759>

Abstract

This study examines students' revision behaviors in AI-feedback-supported environments by analyzing domain- and subskill-level changes in writing, patterns of gains and losses, effort allocation across revisions, and the effectiveness of a process-based metric, the Revision Responsiveness Index. Grounded in cognitive load theory, the study evaluates how reductions in cognitive effort during writing, enabled by artificial intelligence tools, may diminish deeper engagement and learning. A quantitative, within-subjects pretest-posttest design was employed to examine revision outcomes as a writing process. The sample included 1,348 high school students who produced more than 3,800 writing samples across persuasive, informative, and narrative genres. Data were collected in a classroom-based setting using a baseline writing task completed without assistance, followed by a timed revision task supported by structured AI feedback. Revision outcomes were analyzed using the Revision Responsiveness Index (RRI), defined as the change in analytic writing scores between initial and revised drafts, supplemented by domain-level scoring and linguistic analyses. Revision outcomes demonstrated high variability with 48 percent of students improving, 34 percent declining, and 18 percent showed no meaningful change. Students who improved exhibited expected gains across multiple writing domains, while those who declined showed larger and uneven losses. These findings suggest that revision is not inherently beneficial and that ineffective integration of feedback can result in a reduction in overall writing quality.

Keywords: Revision Impact Modeling, Analytic Rubric Scoring, Feedback-Driven Revision, Writing Metrics, Cognitive Load, AI-Mediated Feedback

Introduction

The rapid integration of artificial intelligence into educational settings has transformed how students engage with writing tasks. AI tools now offer the ability to generate text, summarize information, and provide immediate feedback, fundamentally altering the cognitive demands of writing (Deep & Chen, 2025, Xue & Song, 2025). While these tools hold promise for improving efficiency and access, they also raise important questions about the relationship between effort and learning. Research suggests that reducing the cognitive demands of writing may come at a cost (Stadler et al., 2024). Emerging research suggests that when students rely on AI to generate or refine text, they may engage less deeply with the underlying ideas and processes required for learning reducing critical thinking, creativity, and authentic writing skills (Budiyono et al., 2025). Writing is a cognitively intensive activity that requires the organization, evaluation, and synthesis of ideas. As a result, the extent to which students actively engage in these processes is critical to their development as writers and critical thinkers. How AI influences this process is yet to be determined.

Background

Revision represents an important stage in the writing process (Bamberg & Clark, 2019). Unlike initial drafting, which reflects a student's current level of understanding, revision requires students to evaluate their own work, interpret feedback, and make deliberate changes to improve clarity, coherence, and effectiveness (Olive, 2022). This process places significant demands on working memory and requires the coordination of multiple dimensions of writing simultaneously (Fidalgo et al., 2009). Further, since writing demands coordinating multiple interdependent domains, including idea development, organization, language use, and adherence to conventions, effective revision requires students to balance these domains under conditions of limited cognitive capacity (Ahmed et al., 2022).

Despite revision's importance, it does not always lead to overall writing improvement due to multiple factors. Primary obstacles to improving writing through revisions include a combination of superficial focus (Aben et al., 2026); ineffective feedback (Stuulen et al., 2024), emotional and cognitive barriers (Surat et al., 2014), lack of training (Arias-Gundin & Garcia-Sanchez, 2008), and individual differences. Understanding why this occurs requires a closer examination of how students engage with feedback and how they allocate cognitive effort during the revision process.

Cognitive Demands of Writing and Revision

Revision places substantial demands on working memory as students evaluate existing text while attempting to generate improvements across multiple dimensions of writing. Writing is inherently cognitively demanding and requires simultaneously managing multiple processes, including generating ideas, organizing information, selecting appropriate language, and maintaining grammatical accuracy (Fitzgerald, 1987). These demands place significant strain on working memory, particularly for developing writers (Vanderberg & Lee, 2007). Revision further increases cognitive stresses by requiring students to evaluate existing text while simultaneously generating improvements. This dual demand can lead to cognitive overload, particularly when students must address multiple aspects of writing at once (Olive, 2022). Under such conditions, learners may simplify the task by focusing on one dimension of writing while neglecting others.

Writing as a Recursive and Multi-Dimensional Process

Models of writing emphasize its recursive and multi-dimensional nature (Fitzgerald, 1987; Flower & Hayes, 1981). Rather than progressing linearly from planning to drafting to revision, writers continually move between these processes, revising and refining their work as new ideas emerge. Further, effective writing depends on the integration of multiple domains, including content development, organization, language use, and conventions (Ahmed et al., 2022). These domains are interdependent. Adding new ideas may require reorganization, while revising sentence structure may affect clarity and coherence. Revision, therefore, is not simply a matter of correcting errors or adding detail. It requires metacognition, critical thinking, and restructuring ideas in a way that maintains balance across multiple dimensions of writing. This makes revision cognitively complex and can significantly increase cognitive load.

AI and Cognitive Engagement in Writing

The emergence of AI tools in writing instruction has introduced new dynamics into the writing process. While AI can reduce the effort required to generate and refine text, recent research suggests that this reduction in effort may also reduce cognitive engagement. Fan et al. (2025) reported that students who used AI for revision achieved higher scores but demonstrated lower levels of metacognitive engagement, often relying on direct incorporation of AI-generated text. Dhawan et al. (2025) confirmed these dependencies and added that reliance on AI was associated with reduced neural activity and lower retention of content. Stadler et al. (2024) reported that students using AI tools to support writing tasks produced less accurate and less nuanced responses, despite perceiving the task as easier. These findings suggest that AI can shift the cognitive demands of writing in ways that may undermine learning.

Much of this research focuses on environments in which AI contributes directly to content generation or revision. Less is known about how students engage with feedback in environments where AI provides guidance but does not perform the writing itself (Moni, 2025; Wahyuni et al., 2026; Zhan & Yan, 2025). In such contexts, the cognitive demands of revision remain with the learner, creating an opportunity to examine how students interpret and apply feedback.

Revision as a Process of Cognitive Coordination

While cognitive load theory (CLT) and writing models provide insight into the complexity of writing, limited research has examined how these dynamics operate specifically during revision. Revision requires students to interpret feedback, prioritize changes, and apply those changes across multiple domains of writing within a limited time frame (Aben et al., 2026; Fitzgerald, 1987). This process can be understood as a form of cognitive coordination, in which students must allocate attention and effort across competing demands (Ahmed et al., 2022). When this coordination is successful, students achieve balanced improvements across domains. When it is not, revision may result in uneven outcomes, including gains in some areas and declines in others. Understanding revision as a process of cognitive coordination provides a foundation for examining not only whether students improve, but how they improve and where breakdowns occur.

Purpose of Study

Although prior research has examined the effects of AI-supported feedback on writing performance and cognitive engagement, comparatively little attention has been given to how students independently process and apply feedback during revision. The purpose of this study is to examine students' revision behaviors in AI-feedback-supported environments by analyzing domain- and subskill-level changes in writing, patterns of gains and losses, effort allocation across revisions, and the effectiveness of a process-based metric, the Revision Responsiveness Index.

Context is central to this investigation, as most existing studies examine feedback use within instructional or guided learning environments. In contrast, this study focuses on revision behavior when feedback is processed

independently, outside of the teaching environment. Further, this study compares these independent behaviors to feedback engagement within instructional settings. Through this focus, the study seeks to advance understanding of revision as a cognitively demanding process and to inform assessment approaches that more accurately capture how students engage with feedback in AI-supported learning contexts. Theoretically, distinguishing between in-class and independent revision addresses a gap in cognitive and writing research by positioning revision as a context-sensitive process of cognitive coordination, rather than a uniform response to feedback, with implications for how revision outcomes are interpreted and assessed.

Research Questions

- RQ1: How does student writing change between initial and revised drafts after receiving structured, rubric-aligned AI-supported feedback?
- RQ2: What patterns of gains and losses occur across writing domains and subskills during revision?
- RQ3: How do students allocate cognitive effort during revision, and how does this allocation relate to successful versus unsuccessful revision outcomes?
- RQ4: To what extent does the Revision Responsiveness Index (RRI) function as an effective process-based measure of revision impact and cognitive engagement?

Theoretical Framework

To move from descriptive accounts of revision difficulty to explanatory theory, this section applies cognitive load theory (CLT) to clarify how working memory constraints shape students' coordination of multiple writing demands during revision. CLT is applied here as a complementary explanatory framework rather than a causal account of observed revision patterns. CLT provides a foundational framework for understanding how learners process complex tasks (Sweller, 1988; Sweller et al., 2011). CLT asserts that working memory has limited capacity, and learning depends on how effectively cognitive resources are allocated among different types of load: intrinsic, extraneous, and germane. Revision intensifies cognitive load by requiring writers to manage multiple, overlapping cognitive processes simultaneously, including critical evaluation, error detection, and idea refinement (Roussey & Piolat, 2005).

Revision as Cognitive Coordination

Revision plays a central role in writing development, yet it remains undertheorized as a cognitive process through which learners translate feedback into change (Graham & Harris, 2018). Much of the writing literature emphasizes drafting processes or final product quality, offering limited explanation for how writers evaluate feedback, prioritize revisions, and coordinate changes across multiple dimensions of text (Flower & Hayes, 1981). Unlike drafting, which reflects a writer's current understanding, revision reveals how that understanding is actively restructured, making it one of the most cognitively demanding phases of writing (Fidalgo et al., 2009; Olive, 2022).

From a cognitive perspective, revision requires the simultaneous activation of multiple processes, including monitoring existing text, interpreting feedback, generating alternative formulations, and maintaining coherence across the composition (Fidalgo et al., 2009). Writers must hold both the original and intended versions of a text in working memory while making targeted changes, a dual representation that places substantial demands on attentional resources and introduces the potential for instability in performance (Ahmed et al., 2022).

Building on cognitive load theory and models of writing, revision is conceptualized in this study as a process of cognitive coordination, in which writers must allocate limited cognitive resources across multiple, interdependent domains, including idea development, organization, language use, and conventions. Because these domains are interconnected, changes in one area can produce unintended effects in others. For example, expanding an argument may improve idea development while introducing organizational inconsistencies or grammatical errors, and revising sentence structure for clarity may disrupt tone or stylistic coherence.

Effective revision therefore depends on the writer's ability to negotiate competing demands under cognitive constraint. When this coordination is successful, revision results in balanced improvement across domains. When coordination falters, revision may produce uneven outcomes, including gains in some areas alongside declines in others. These challenges are amplified when writers must address both lower-level revisions, such as grammar and spelling, and higher-level revisions, such as coherence and argumentation, within limited time and cognitive capacity (Roussey & Piolat, 2005). Less skilled writers tend to revise more frequently but at smaller scopes, a pattern that can further increase cognitive load and undermine text quality. Emotional factors, including frustration or reduced motivation during revision, may further compound these demands.

Feedback, Decision-Making, and Revision Behavior

Feedback plays a critical role in guiding revision, but its effectiveness depends on how the writer interprets and applies it (Hattie & Timperley, 2007). In many instructional contexts, feedback is assumed to lead directly to improvement. However, this assumption overlooks the cognitive work required to translate feedback into meaningful changes (Stadler et al. 2022). Students must first understand the feedback, then decide which elements to prioritize, and finally determine how to implement those changes within the structure of their existing text. Each of these steps introduces potential points of breakdown. Feedback may be immediate and detailed in AI-supported environments, but it does not reduce the need for student decision-making. Instead, it may increase cognitive demands by presenting multiple areas for improvement simultaneously (Roodsari et al., 2024). Without the ability to prioritize and coordinate changes, students may focus narrowly on specific aspects of feedback, leading to fragmented revision. Feedback alone is insufficient to support learning. The critical factor is how students engage with feedback during revision.

Toward a Process-Based Measure of Revision

Traditional approaches to writing assessment have predominantly emphasized final product quality, offering limited insight into the cognitive and strategic processes through which revisions are generated (Flower & Hayes, 1981). While such measures provide useful information about performance, they offer limited insight into the processes that produce that performance (Graham & Harris, 2018). A process-based approach to assessment shifts attention from outcomes to writing growth. Evaluating how writing evolves between drafts makes it possible to capture nuances of learning that are not visible in a single snapshot of performance. Recognizing these nuances transforms the writing process into a measurable indicator of cognitive engagement. Iterations in writing reflect how students interpret feedback, allocate attention, and manage competing demands.

Methodology & Design

This study employed a within-subjects design to examine changes in student writing following structured feedback and revision. Each student completed an initial writing task, received targeted feedback aligned to a standardized rubric, and then revised their writing within a constrained time frame. The design allowed for direct comparison between initial and revised drafts, enabling analysis of revision impact at the overall, domain, and subskill levels. The study was conducted in authentic classroom settings, reflecting typical instructional conditions while maintaining consistency in task structure and scoring procedures.

Participants and Data Set

The sample included 1,348 high school students from one Mid-Atlantic United States public school district across multiple course levels, including advanced, academic, and lower-track classes. Students produced more than 3,800 writing samples across three genres: persuasive, informative, and narrative. This large and diverse sample provides a robust basis for examining patterns of revision behavior across varying levels of student proficiency and instructional contexts.

Writing Tasks and Conditions

Students completed a baseline writing task (writing on demand) in response to a standardized prompt aligned to one of three genres: persuasive, informative, or narrative. All writing was completed independently without access to external resources or generative AI tools. Following the initial draft, students received structured feedback aligned to the analytic rubric used for scoring. Feedback was provided at the domain and subskill levels, offering specific guidance on areas for improvement.

Students then engaged in a timed revision period of approximately 15 minutes. During this time, they were required to interpret and apply feedback independently. No generative AI tools or external assistance were permitted during revision. This condition ensured that any changes in writing reflected student cognitive engagement rather than automated text generation.

Scoring Framework

Each writing sample was evaluated using a structured analytic rubric consisting of five domains. Four domains were consistent across all genres: Ideas and Purpose; Organization; Language and Style; and Conventions and Knowledge. The fifth domain was genre-specific: Narrative Technique for narrative writing and Evidence and Development for persuasive and informative writing. Each domain consisted of five subskills, with each subskill scored on a scale from 0 to 10. This resulted in a maximum of 50 points per domain and a total possible score of 250 points per writing sample. Scoring was conducted consistently across both initial and revised drafts using the same rubric and criteria. This ensured that differences in scores reflected changes in student writing rather than variations in evaluation standards. The use of subskill-level scoring enabled detailed analysis of writing performance, allowing for examination of changes within and across domains.

Conceptualizing the Revision Responsiveness Index (RRI)

The Revision Responsiveness Index (RRI) is introduced as a process-based construct designed to capture how students engage with feedback during revision. Unlike traditional assessment models that focus solely on final product quality, RRI emphasizes the process of change between drafts. In its broader conceptual model, RRI is defined as a multi-dimensional construct encompassing three key components: Responsiveness to Feedback (the degree to which students address targeted feedback; Revision Efficiency (the effectiveness of changes made within a constrained time frame); and Impact on Writing Quality (the measurable change in writing performance across domains). These dimensions define the broader conceptual scope of RRI as an index of students' cognitive engagement during revision. In the present study, RRI is operationalized specifically as a measure of revision impact, defined as the change in analytic writing scores between initial and revised drafts. While responsiveness and efficiency are central to the broader conceptual model, they are not directly measured in this analysis and are treated as areas for future research.

RRI as a Measure of Revision Impact

Within the context of this study, RRI captures the extent to which student writing improves or declines following feedback. This approach shifts the focus of assessment from static performance to dynamic change. By examining differences between initial and revised drafts, RRI provides insight into how students process and apply feedback. Positive RRI values indicate improvement, while negative values indicate decline. Importantly, this formulation allows for the identification of factors that lead to the variability in revision outcomes, including cases where revision does not lead to improvement. This perspective aligns with a process-based view of learning, in which growth is reflected not only in final outcomes but in the ability to respond effectively to feedback.

Operational Definition of the Revision Responsiveness Index (RRI)

The Revision Responsiveness Index (RRI) is operationalized in this study as the difference between a student's total writing score on a revised draft and their total writing score on an initial draft, based on a structured analytic rubric. RRI as operationalized is a study-specific indicator of revision impact, not a validated comprehensive assessment instrument. Each writing sample is evaluated across five domains. Four domains are consistent across genres: Ideas and Purpose, Organization, Language and Style, and Conventions and Knowledge. The fifth domain is genre-specific, with Narrative Technique used for narrative writing and Evidence and Development used for informative and persuasive writing. Each domain consists of five subskills, each scored on a scale from 0 to 10. Subskill scores are aggregated to produce a maximum of 50 points per domain and a total possible score of 250 points per writing sample. RRI is calculated as: $RRI = \text{Final Total Score} - \text{Initial Total Score}$.

The Final Total Score represents the sum of all domain and subskill scores in the revised draft, and Initial Total Score represents the sum of all domain and subskill scores in the initial draft. Expanded, this calculation can be expressed as: $RRI = \sum (D_{post} - D_{pre})$ where D_{pre} represents the score for each subskill within each domain in the initial draft, and D_{post} represents the corresponding score in the revised draft. Domain-level revision is calculated as: $\text{Domain Delta}_i = \text{Domain Post}_i - \text{Domain Pre}_i$. Positive RRI and Domain Delta values indicate net gains across drafts, while negative values indicate net losses. This formulation enables analysis at multiple levels, including overall writing change, domain-specific gains and losses, and subskill-level variation. In this study, RRI reflects the impact dimension of revision. Future iterations of the model may incorporate additional measures of responsiveness to feedback and revision efficiency to support a more comprehensive process-based index.

Data Analysis

Revision was measured as the change in writing performance between the initial and revised drafts. Differences were calculated at three levels: Overall score change (total score out of 250); Domain-level change (score out of 50 per domain); and Subskill-level change (score out of 10 per subskill). These differences form the basis for calculating the Revision Responsiveness Index (RRI), as defined in Section 4, and for analyzing patterns of revision behavior across the sample. Descriptive analyses were conducted to examine overall trends in revision outcomes, including the proportion of students who improved, declined, or remained unchanged. Domain-level and subskill-level changes were analyzed to identify patterns of growth and decline across different aspects of writing. In addition, quantitative writing metrics, including word count, lexical diversity, and sentence-level measures, were analyzed to examine changes in writing characteristics associated with revision.

Ethical Considerations

All data were collected in accordance with standard educational research practices. Student data were anonymized prior to analysis to ensure privacy and confidentiality. The study was conducted within normal instructional activities, and no personally identifiable information was used in reporting results.

Results

Differences between initial and revised draft scores were analyzed across 1,348 students and more than 3,800 writing samples. A paired-samples t-test indicated a statistically significant difference between baseline and revised scores, $t(3875) = -21.41, p < .001$, with a moderate effect size (Cohen's $d = .34$; 95% CI $[-.376, -.311]$). Mean scores increased from 142.8 to 150.1, reflecting an average gain of 7.3 points on the 250-point scale. This indicates that, on average, revision resulted in measurable improvement in writing performance. Across the sample, approximately 48% of students demonstrated improvement in total writing score, 34% experienced a decline, and 18% showed no meaningful change (see Figure 1).

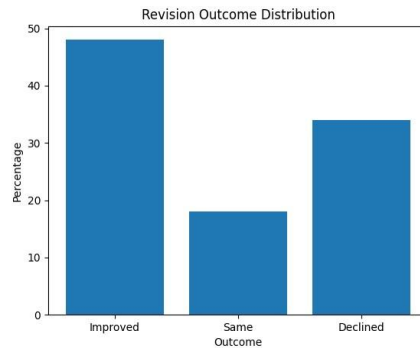


Figure 1: *Distribution of Revision Outcomes.*

Note: The percentage of students whose writing improved, declined, or showed no meaningful change following revision.

Overall Revision Responsiveness Index (RRI)

The Revision Responsiveness Index (RRI), defined as the difference between revised and initial total writing scores, provides a measure of revision impact across the analytic scoring framework. Across the full sample, the mean RRI was 7.37 on a 250-point scale, closely aligning with the observed composite gain of 7.3 points derived from paired comparisons. At the aggregate level, the observed effect size (Cohen's $d = .34$) reflects a modest but meaningful level of improvement within a single revision cycle. RRI values varied widely across students.

Domain-Level Changes Across the Full Sample

Domain-level analysis revealed statistically significant improvements across all writing dimensions when comparing initial and revised drafts ($p < .001$). Mean gains were observed in each domain, indicating that revision produced measurable changes across multiple aspects of writing performance. Average domain gains (on a 50-point scale) were as follows:

- Organization: +1.72
- Evidence & Development / Narrative Technique: +1.69
- Narrative Technique / Evidence & Development: +1.52
- Ideas & Purpose: +1.46
- Language & Style: +1.42
- Conventions & Knowledge: +1.06

These gains correspond to moderate and consistent effect sizes across higher-order domains (Cohen's $d \approx .32-.34$), with smaller but still meaningful gains observed in Conventions & Knowledge ($d = .20$).

Linguistic Characteristics of Revision

To better understand how students revise, quantitative writing metrics were analyzed, including word count, lexical diversity, and sentence-level measures. On average, students increased word count by 28.62 words and demonstrated gains in lexical diversity (+2.9). In contrast, words per sentence showed minimal change (−0.07), indicating little to no increase in syntactic complexity (see Figure 2).

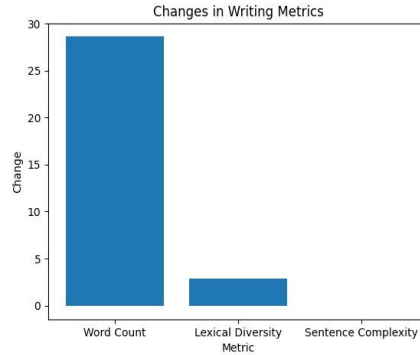


Figure 2: *Linguistic Changes Following Revision*

Note: Average changes in word count, lexical diversity, and sentence-level complexity between initial and revised drafts.

Domain-Level Decline Patterns

To better understand unsuccessful revision, domain-level changes were examined for students whose total writing scores declined following revision. Students in this group exhibited substantial and uneven losses across domains. The largest average declines were observed in:

- Conventions & Knowledge: -3.22
- Evidence & Development: -2.83
- Narrative Technique: -2.75
- Ideas & Purpose: -2.13
- Language & Style: -1.72
- Organization: -1.69

As shown in Figure 3, declines were most pronounced in foundational and evidence related domains. Notably, losses in Conventions & Knowledge were larger in magnitude than gains observed in the full sample

Domain-Level Gains for Improving Students

In contrast, students whose total writing scores increased demonstrated substantial and coordinated gains across domains. Average domain gains for this group were as follows:

- Ideas & Purpose: $+5.10$
- Evidence & Development: $+3.66$
- Conventions & Knowledge: $+3.14$
- Organization: $+2.75$
- Language & Style: $+2.47$
- Narrative Technique: $+2.31$

These gains, illustrated in Figure 3, are markedly larger than the average domain changes observed in the full sample.

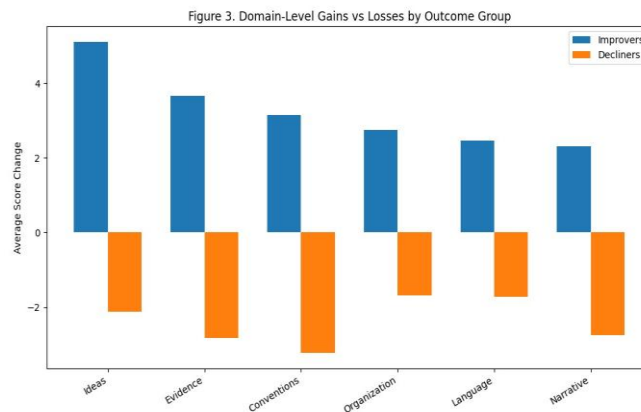


Figure 3: *Domain-Level Gains vs Losses by Outcome Group.*

Note: Average domain score changes for students who improved compared to those who declined following revision.

Cross-Domain Trade-Off Patterns

Comparison of students who improved with those who declined reveals a consistent pattern of cross-domain trade-offs in revision behavior (see Figure 4).

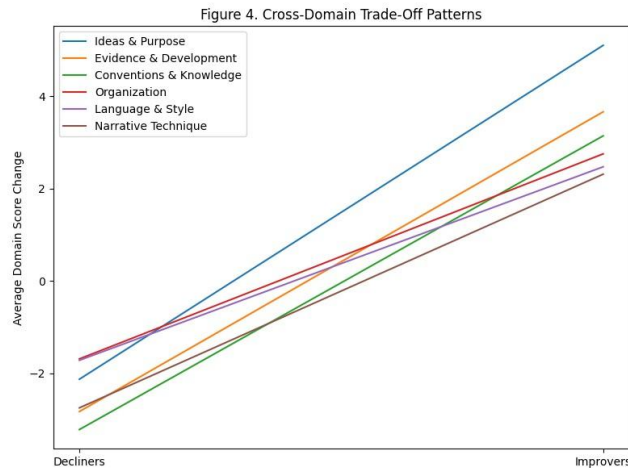


Figure 4: *Cross-Domain Trade-Off Patterns in Revision Outcomes.*

Note: Average domain-level score changes for students who declined compared to those who improved following revision. Each line represents a writing domain, illustrating asymmetry between gains and losses across outcome groups.

Students who demonstrated overall improvement exhibited coordinated gains across multiple domains, particularly in higher-order areas such as Ideas & Purpose and Evidence & Development. These gains were accompanied by stable or improving performance in lower-level domains. In contrast, declining students showed larger and more uneven domain-level losses compared to gains observed in improving students, with particularly pronounced decreases in Conventions & Knowledge and evidence-related domains. These losses were not isolated but distributed across multiple areas of writing.

Relationship Between RRI and Domain-Level Change

Analysis of RRI in relation to domain-level changes indicates that overall revision impact reflects the cumulative effect of gains and losses across domains. Students with higher RRI scores consistently demonstrated positive changes across multiple domains, while lower or negative RRI scores were associated with uneven or fragmented revision patterns.

Summary of Results

Revision produced statistically significant but modest average gains in writing scores; however, outcomes were highly variable across students. Domain-level analyses revealed divergent patterns, with coordinated multi-domain gains among improving students and larger, uneven losses among declining students. Linguistic analyses showed increased text length and lexical diversity with minimal change in syntactic structure. These findings indicate substantial heterogeneity in revision outcomes and highlight the limits of average score change in capturing revision behavior.

Discussion

Across the sample, revision outcomes were highly variable, with fewer than half of students demonstrating net improvement and a substantial proportion experiencing declines or no meaningful change. Although average gains in total writing scores were statistically significant, deeper analyses revealed that these aggregate improvements masked divergent revision trajectories. Mean score changes concealed substantial differences in how students distributed gains and losses across writing domains, underscoring the limitations of relying solely on average outcomes to characterize revision behavior.

Domain-level analyses indicated that students who improved exhibited coordinated gains across multiple dimensions of writing, particularly in higher-order domains such as Ideas and Purpose and Evidence and Development, without corresponding declines in lower-level features. In contrast, students whose writing declined showed larger and uneven losses distributed across domains, with particularly pronounced decreases in Conventions and Knowledge and evidence-related domains. In many cases, the magnitude of these losses exceeded the average domain gains observed across the full sample, suggesting that unsuccessful revision can

introduce substantial degradation rather than simply lacking improvement. Linguistic analyses further indicate that revision was typically characterized by additive strategies, such as increasing text length and lexical diversity, rather than structural transformation, potentially compounding cognitive demands when revisions were not effectively coordinated.

Taken together, these patterns suggest that revision is not a neutral or uniformly beneficial process, but one that involves the redistribution of cognitive effort across competing writing demands. Successful revision appears to depend on students' ability to integrate feedback across domains in a balanced manner, whereas unsuccessful revision is marked by fragmented or misaligned application of changes. The Revision Responsiveness Index captures this variability by reflecting the cumulative impact of domain-level gains and losses, providing a process-sensitive measure of revision outcomes that is not apparent in aggregate score changes alone.

Reframing Revision as Cognitive Coordination

To interpret these empirical patterns within existing theory, revision can be reframed as a process of cognitive coordination. The findings of this study align with the literature finding that writing is recursive, not linear (Fidalgo et al., 2009), and is a cognitively demanding act of coordination across multiple dimensions of writing (Vanderberg & Lee Swanson, 2007). Students are required to simultaneously manage ideas, evidence, organization, language, and conventions within the constraints of working memory. Under these conditions, revision becomes a process of prioritization and trade-off rather than uniform enhancement.

The observed asymmetry between gains and losses reinforces this interpretation. Students who improved demonstrated coordinated, multi-domain growth, while those who declined exhibited larger, fragmented losses across domains. This pattern indicates that successful revision depends on the ability to integrate feedback holistically, whereas unsuccessful revision reflects misaligned or isolated changes.

This conceptualization positions revision as a form of cognitive coordination, in which students must allocate limited mental resources across competing demands. When coordination is successful, writing improves across domains. When it is not, revision efforts may result in unintended degradation of writing quality.

Alignment with Cognitive Load Theory

These findings align closely with CLT, which posits that learning depends on the effective management of intrinsic, extraneous, and germane cognitive load (Sweller, 1988; Sweller et al., 2011). Writing and revision inherently involve high intrinsic load due to the complexity of coordinating multiple skills simultaneously. The domain-level trade-offs observed in this study suggest that cognitive demands may exceed some students' available processing capacity. As cognitive demands intensify, revision efforts may emphasize one domain while other dimensions receive reduced attention. For example, efforts to strengthen ideas or evidence may lead to declines in conventions or organization, reflecting a redistribution rather than an expansion of cognitive effort. These findings extend cognitive load theory into the domain of writing revision by demonstrating how overload manifests in measurable patterns of domain-level gains and losses. Rather than improving uniformly, students exhibit uneven development that reflects the constraints of working memory.

Implications for AI-Supported Writing Environments

The results of this study provide important context for understanding the role of AI in writing instruction. Recent research has shown that generative AI tools can reduce cognitive effort while also diminishing depth of understanding and metacognitive engagement (Stadler et al., 2024; Fan et al., 2025). This study's findings raise concerns about the potential for AI to bypass the very processes that support learning. In this respect, AI integration in writing instruction requires careful consideration of how cognitive effort is distributed rather than simply whether performance improves.

This study extends this line of inquiry by demonstrating that even when students are required to revise independently, without access to AI-generated text, effective revision outcomes were associated with active student engagement during feedback use. Feedback alone did not guarantee improvement across drafts. Instead, students were required to interpret, prioritize, and apply feedback in a coordinated manner. These findings indicate that cognitive engagement remains a critical determinant of revision success, even in the absence of AI-generated writing support. This distinction has important implications for how AI is implemented in writing instruction. When AI is used to generate or rewrite text, it may remove the need for students to engage in cognitive coordination altogether. Such use may obscure underlying revision challenges rather than resolve them. In contrast, when AI is used to provide feedback that must be actively processed, it preserves the cognitive demands necessary for learning. These findings suggest that the instructional value of AI depends more on how it is integrated into the learning process than on its availability.

Revision as the Site of Learning

A central implication of this study is that revision represents the primary site of learning in writing. While initial drafting reflects students' current understanding, revision reveals how they engage with feedback and whether they can transform that understanding.

The variability observed in revision outcomes indicates that learning does not occur automatically through the feedback. Instead, it depends on the quality of cognitive engagement during revision. Students who demonstrated meaningful growth were able to apply feedback across domains in a coordinated manner, while those who declined often applied feedback inconsistently. This reinforces the idea that writing instruction should focus on how they revise, not only the end product. The process of revision, rather than the final product alone, provides a more accurate window into student learning.

The Role of the Revision Responsiveness Index (RRI)

The Revision Responsiveness Index provides a framework for operationalizing revision as a measurable construct. By capturing responsiveness to feedback, efficiency of change, and impact on writing quality, RRI offers a process-based indicator of student learning. Importantly, RRI does not simply reflect improvement in final scores. Instead, it captures the degree to which students engage in effective revision behaviors. High RRI scores are associated with coordinated, multi-domain gains, while lower scores reflect fragmented or ineffective revision patterns. In this way, RRI serves as a bridge between theoretical constructs such as cognitive load and metacognition and observable classroom outcomes. It enables educators to identify not only whether students improve, but how they improve, and where breakdowns occur.

Instructional Implications

The findings of this study have several implications for instructional practice in writing and AI-supported learning environments. First, students require explicit instruction in how to revise. Feedback alone is insufficient; students must be taught how to interpret feedback, prioritize revisions, and coordinate changes across domains. Second, revision should be structured as a bounded and intentional process. Time constraints, such as a 15-minute revision window, may encourage students to engage more actively in decision-making and reduce reliance on passive or surface-level strategies. Third, instructional models should emphasize the interconnected nature of writing domains. Students need to understand that changes in one area of writing can affect others, and that effective revision requires maintaining balance across domains. Finally, AI should be used to support, rather than replace, cognitive engagement. Tools that provide feedback without generating text may be more effective in promoting learning than those that automate the writing process.

Summary

This study advances a new understanding of revision as a cognitively constrained coordination process. Rather than functioning as a straightforward mechanism for improvement, revision represents a complex interaction of competing demands that can produce either meaningful growth or unintended decline. In AI-supported environments, this distinction becomes especially important. Technologies that reduce cognitive effort may inadvertently bypass the processes that support learning, while those that preserve cognitive engagement can enhance it. By shifting the focus from product to process, and from output to cognitive engagement, this work contributes to a broader redefinition of writing instruction and assessment. The Revision Responsiveness Index offers a practical and theoretically grounded tool for capturing these processes, providing educators with new insight into how students learn through writing.

Limitations and Directions for Future Research

The study's quasi-experimental design in real classroom settings increases ecological validity but also introduces contextual variability that limits causal inference and reduces generalizability beyond secondary students or other instructional models. The Revision Responsiveness Index (RRI), while a practical for cognitive engagement, infers internal processes rather than directly measuring them, requiring further validation of its reliability, sensitivity, and alignment with established engagement and metacognitive measures, especially across longer writing trajectories. Additionally, because generative AI use was intentionally restricted, future research should experimentally compare AI-assisted, feedback-only, and independent revision; examine revision development across multiple writing cycles; and conduct domain-level analyses to determine which writing subskills are most responsive to revision and whether targeted interventions. Explicit instruction in feedback interpretation and revision planning in future studies could strengthen revision outcomes.

Conclusion

The integration of generative artificial intelligence into writing instruction represents one of the most significant shifts in modern education. While these tools offer unprecedented efficiency and access, they also challenge long-

standing assumptions about how writing supports learning. In environments where high-quality text can be produced with minimal effort, the traditional link between writing performance and cognitive engagement is no longer guaranteed. This study responds to that challenge by shifting the focus of assessment from product to process. Rather than asking what students produce, it asks how students think and how they engage with feedback and revise their work.

The findings demonstrate that revision is neither automatic nor uniformly beneficial. Even within a structured, feedback-rich environment, student outcomes varied widely. While nearly half of students improved, a substantial proportion experienced no change or decline. These results underscore a critical insight: revision is a complex cognitive activity that can either advance or undermine learning depending on how it is executed.

Patterns observed in the data further reinforce this conclusion. Students tended to increase word count and lexical diversity yet showed limited evidence of deeper structural change. Improvements in one domain were frequently accompanied by declines in another, revealing the complexity of coordinating multiple dimensions of writing. In many cases, students believed they had improved because they followed feedback, yet their performance declined due to fragmented or misaligned revision strategies. These findings challenge the assumption that access to feedback or tools is sufficient to support learning. Instead, they point to the importance of cognitive engagement as the central driver of writing development. Learning occurs when feedback is processed, interpreted, and meaningfully applied.

The Revision Responsiveness Index (RRI) addresses this gap in the revision process. By capturing responsiveness to feedback, efficiency of engagement, and impact on writing quality, RRI provides a practical and scalable way to observe the learning process. It offers educators a means of distinguishing between superficial revision and cognitively engaged improvement, bridging the gap between instructional practice and assessment.

This distinction is essential in the context of AI-supported writing. The question is no longer whether students can produce strong writing, but whether they are engaging in the thinking required to develop as writers. Tools that generate text may accelerate production and appear to develop writing skills, but they do not inherently support learning unless they are embedded within systems that require active cognitive participation.

This study suggests that writing instruction must not neglect the revision and feedback process and proficiency and advancement does not depend merely on current measurements. Educators should design learning that encourages thinking by emphasizing feedback as a tool for reflection. Teaching writing should include structured revision as an active cognitive task and should incorporate process-based metrics that make this process visible. By reframing writing assessments around cognitive engagement, educators can better align instruction with the realities of AI and ensure that writing continues to function as a vehicle for learning.

Writing remains one of the most powerful tools for developing human thought. Removing technology from writing will not solve this problem. The challenge in the age of artificial intelligence is to ensure that technology is used in ways that preserve and expand the cognitive work that writing demands. The Revision Responsiveness Index represents a step in that direction, offering a new framework for understanding, measuring, and supporting learning through writing in an increasingly AI mediated world.

References

- Aben, J. E. J., de Jong, J., Coupé, G., Meijntz, M., & Bouwer, R. (2026). Who's talking? Dialogic peer feedback shifts students' focus from superficial to meaningful revisions. *Reading and Writing*, 39, 753–776. <https://doi.org/10.1007/s11145-025-10743-8>
- Ahmed, Y., Kent, S., Cirino, P. T., & Keller-Margulis, M. (2022). The not-so-simple view of writing in struggling readers/writers. *Reading & Writing Quarterly*, 38(3), 272-296. DOI: 10.1080/10573569.2021.1948374
- Arias-Gundín, O., & García-Sánchez, J. N. (2008). Educational implications of the theoretical models of text revision. *Papeles del Psicólogo*, 29(2), 222-228.
- Bamberg, B., & Clark, I. L. (2019). Revision: Issues and strategies. In I. L. Clark (Ed.), *Concepts in composition: Theory and practices in the teaching of writing* (3rd ed.). Routledge. <https://doi.org/10.4324/9780203728659-3>
- Budiyono, H., Marzuki, M., Pudjaningsih, W., Prastio, B., & Maulidina, A. (2025). Exploring long-term impact of AI writing tools on independent writing skills: A case study of Indonesian language education students. *International Journal of Information and Education Technology*, 15(5), 1003-1013. DOI 10.18178/ijiet.2025.15.5.2306

- Deep, P. D., & Chen, Y. (2025). The role of AI in academic writing: Impacts on writing skills, critical thinking, and integrity in higher education. *Societies*, 15(9), 247. 10.3390/soc15090247
- Dhawan, N., Bhasin, S., Gupta, A., & Khalkho, J. T. (2025). Understanding the impact of AI on the cognitive thinking of students. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.5433395>
- Fitzgerald, J. (1987). Research on revision in writing. *Review of educational research*, 57(4), 481-506. DOI: 10.3102/00346543057004481
- Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., et al. (2024). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*. Advance online publication. <https://doi.org/10.1111/bjet.13429> (Preprint available: <https://arxiv.org/abs/2412.09315>)
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387. <https://doi.org/10.2307/356600>
- Graham, S., & Harris, K. R. (2018). Evidence-based writing practices: A meta-analysis. *Reading and Writing*, 31, 1333–1360. DOI: <https://doi.org/10.1007/s11145-018-9848-2>
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112. <https://www.proquest.com/scholarly-journals/power-feedback/docview/214113991/se-2>
- Olive, T. (2022). *How do novice and skilled writers engage working memory?* The Cambridge handbook of working memory and language, 504-526. DOI: 10.1017/9781108955638.028
- Roodsari, S. T., Köhler, T., & Azizi Ghanbari, S. (2024, September). *Artificial Intelligence-based Learning Objective-Oriented Formative Feedback for Students in Higher Education Based on the Instructional Design*. In International Conference on Interactive Collaborative Learning (pp. 504-513). Cham: Springer Nature Switzerland. DOI: 10.1007/978-3-031-85652-5_50.
- Roussey, J. Y., & Piolat, A. (2005). La révision du texte: une activité de contrôle et de réflexion. *Psychologie française*, 50(3), 351-372. DOI: 10.1016/j.psfr.2005.05.001
- Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, 160, 108386. <https://doi.org/10.1016/j.chb.2024.108386>
- Stuulen, J., Bouwer, R., & van den Bergh, H. (2024). Effects of a comparative feedback method on peer feedback characteristics and revision quality. *L1-Educational Studies in Language and Literature*, 24(1), 1-21. <https://doi.org/10.21248/l1esll.2024.24.1.671>
- Surat, S., Rahman, S., Mahamod, Z., & Kummin, S. (2014). The Use of Metacognitive Knowledge in Essay Writing among High School Students. *International Education Studies*, 7(13), 212-218. DOI: 10.5539/ies.v7n13p212
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational psychology review*, 22(2), 123-138. 10.1007/s10648-010-9128-5
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer. <https://doi.org/10.1007/978-1-4419-8126-4>
- Vanderberg, R., & Lee Swanson, H. (2007). Which components of working memory are important in the writing process? *Reading and Writing*, 20(7), 721-752. DOI: 10.1007/s11145-006-9046-6
- Wahyuni S., Dewi D.S., Wulandari R.S., Thohiriyah, Kurniadi D., Muawanah F.H., Chairunnisa N.Z., and Ardhi M.A. (2026). Evaluating students' perceptions, writing outcomes, and challenges in using AI-based feedback systems for English academic writing. *Multidisciplinary Reviews*, 9(8) Article no. e2026380. DOI: 10.31893/multirev.2026380
- Zhan, Y., & Yan, Z. (2025). Students' engagement with ChatGPT feedback: Implications for student feedback literacy in the context of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 1-14. DOI: 10.1080/02602938.2025.2471821
- Xue, R., & Song, Q. (2025, July). *The Impact of Artificial Intelligence-Generated Content on IELTS Part 2 Writing Instruction*. In Proceedings of the 2025 3rd International Conference on Educational Knowledge and Informatization (pp. 1-6). DOI: 10.1145/3765325.3765327